

**DISERTASI
(TEK 010317)**

**IMPUTASI DATA DERET WAKTU MULTIVARIAT
MENGUNAKAN METODE SKDL (STING KERNEL DEEP
LEVEL) WITH EXPLAINABLE**



Nama : Arius Satoni Kurniawansyah
NIM : 03013682126021
BKU : Teknik Informatika
Promotor : Prof. Dr. Ir. Siti Nurmaini, MT
Ko - Promotor : Prof. Dr. Erwin, S.Si., M.Si

**PROGRAM STUDI DOKTOR ILMU TEKNIK
FAKULTAS TEKNIK
UNIVERSITAS SRIWIJAYA
2025**

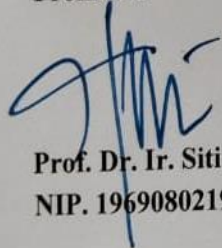
**HALAMAN PENGESAHAN
DISERTASI
(TEK 010317)**

**IMPUTASI DATA DERET WAKTU MULTIVARIAT
MENGUNAKAN METODE SKDL (STING KERNEL DEEP
LEVEL) WITH EXPLAINABLE**

**Oleh:
ARIUS SATONI KURNIAWANSYAH
NIM. 03013682126021**

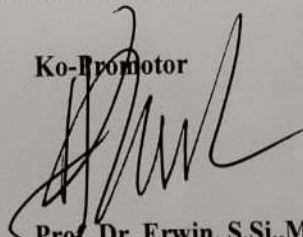
**Telah disetujui
Pada Tanggal 06, Bulan Agustus, Tahun 2025**

Promotor



**Prof. Dr. Ir. Siti Nurmaini, MT
NIP. 196908021994012001**

Ko-Promotor



**Prof. Dr. Erwin, S.Si., M.Si
NIP. 197101291994121001**

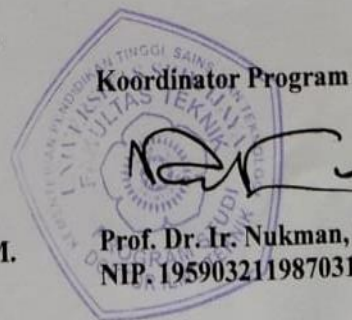
Mengetahui,

Dekan Fakultas Teknik,



**Dr. Ir. Bhakti Yudho Suprpto, S.T., M.T., IPM.
NIP. 197502112003121002**

Koordinator Program Studi



**Prof. Dr. Ir. Nukman, M.T.
NIP. 195903211987031001**

HALAMAN PERSETUJUAN

Disertasi berjudul "Imputasi Data Deret Waktu Multivariat Menggunakan Metode SKDL (STING Kernel Deep Level) With Explainable" telah dipresentasikan dihadapan Tim Penguji Disertasi pada Program Studi Doktor Ilmu Teknik Fakultas Teknik Universitas Sriwijaya pada hari Senin, tanggal 25, bulan Agustus, tahun 2025.

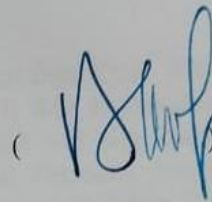
Palembang, tanggal 25, bulan Agustus, tahun 2025

Tim Penguji Disertasi berupa Disertasi:

Ketua Tim Penguji:

Nama : Dr. Ir. Bhakti Yudho Suprpto, S.T., M.T., IPM

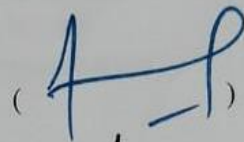
NIP : 197502112003121002

()

Anggota Tim Penguji:

1. Nama : Dr. M. Fachrurrozi, M.Kom

NIP : 198005222008121002

()

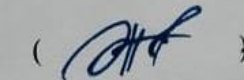
2. Nama : Dr. Anita Desiani, M.Kom

NIP : 197712112003122002

()

3. Nama : Dr. Yesi Novaria Kunang, S.T., M.Kom

NIP : 030302208

()

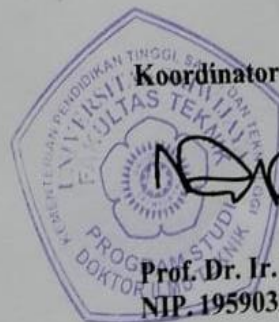
Mengetahui,

Dekan Fakultas Teknik,



Dr. Ir. Bhakti Yudho Suprpto, S.T., M.T., IPM.
NIP. 197502112003121002

Koordinator Program Studi



Prof. Dr. Ir. Nukman, M.T.
NIP. 195903211987031001

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Arius Satoni Kurniawansyah
NIM : 03013682126021
Program Studi : Doktor Ilmu Teknik
BKU : Teknik Informatika

Dengan ini menyatakan bahwa disertasi saya dengan judul “Imputasi Data Deret Waktu Multivariat Menggunakan Metode SKDL (STING Kernel Deep Level) With Explainable”, bebas dari fabrikasi, falsifikasi, plagiat, kepengarangan yang tidak sah dan konflik kepentingan dan pengajuan jamak, seperti yang tercantum dalam Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia Nomor 39 Tahun 2021.

Bilamana ditemukan ketidak sesuaian dengan hal-hal di atas, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan aturan yang berlaku.
Demikian pernyataan ini dibuat dengan sesungguhnya dan dengan sebenar-benarnya.

Palembang, 06 Agustus 2025
Yang menyatakan



Arius Satoni Kurniawansyah
NIM. 03013682126021

ABSTRAK

Imputasi Data Deret Waktu Multivariat Menggunakan Metode SKDL (STING Kernel Deep Level) With Explainable

Penelitian ini bertujuan untuk mengembangkan metode baru dalam mengatasi permasalahan nilai hilang (missing values) pada data deret waktu multivariat (multivariate time series), khususnya pada data tanda vital pasien ICU yang bersumber dari dataset MIMIC-IV. Kehilangan data dalam konteks medis dapat mengganggu proses analisis, prediksi, bahkan pengambilan keputusan klinis. Oleh karena itu, pendekatan imputasi yang akurat dan dapat dijelaskan (explainable) menjadi kebutuhan penting. Penelitian ini memperkenalkan pendekatan gabungan antara dua metode: metode STING (Self-Attention-based Time Series Imputation using GAN) dan metode Kernel Tahap Lanjutan. Metode STING Tahap Lanjutan dirancang untuk menghasilkan imputasi numerik yang presisi melalui mekanisme attention dan arsitektur GAN, sementara metode Kernel difokuskan untuk menghasilkan imputasi yang lebih baik dalam dunia medis. Gabungan keduanya disebut sebagai SKDL (STING Kernel Deep Level with Explainable), yaitu metode imputasi dengan kemampuan dual-output (numerik dan kategorikal) serta dilengkapi komponen Explainable AI. Evaluasi dilakukan menggunakan metrik MAE, MSE, RMSE, R-Squared, serta Sensitivitas, Spesifisitas, dan F1-Score. Hasil menunjukkan bahwa metode STING Tahap Lanjutan menghasilkan MAE 0,090 dan RMSE 0,0109, yang lebih baik dibandingkan metode existing seperti KNN, MICE, dan Mean Imputation. Adapun metode Kernel Tahap Lanjutan mencapai F1-Score 0,825. Metode SKDL mencapai MAE sebesar 0,0870, MSE sebesar 0,0175, RMSE sebesar 0,0040, dan R-Squared sebesar 0,3367, yang menunjukkan hasil dengan akurasi terbaik dibandingkan dengan Metode STING Tahap Lanjutan dan Kernel Tahap Lanjutan. Hasil Penelitian ini menunjukkan bahwa pendekatan yang dikembangkan tidak hanya lebih akurat dalam imputasi, tetapi juga mampu memberikan hasil dengan klasifikasi status klinis pasien secara interpretative (dalam bentuk data kategorikal). Penelitian ini memberikan kontribusi penting dalam bidang data science medis, khususnya pada pengembangan sistem pengambilan keputusan berbasis AI di unit perawatan intensif.

Kata Kunci: Imputasi, Time Series, Multivariat, Data, STING, Kernel, Tahap Lanjutan, Penjelasan

ABSTRACT

Imputation Of Multivariate Time Series Data Using The STING Kernel Deep Level Method With Explainable

This study aims to develop a new method to address the problem of missing values in multivariate time series data, particularly in vital sign data of ICU patients derived from the MIMIC-IV dataset. Missing data in the medical context can hinder analysis, prediction, and even clinical decision-making. Therefore, an accurate and explainable imputation approach is crucial. This research introduces a combined approach involving two methods: the STING method (Self-Attention-based Time Series Imputation using GAN) and the Enhanced Kernel method. The Enhanced STING method is designed to generate precise numerical imputations using attention mechanisms and a GAN architecture, while the Kernel method focuses on achieving better imputations in the medical domain. The combination of these two methods is referred to as SKDL (STING Kernel Deep Level with Explainable), an imputation method capable of producing dual outputs (numerical and categorical) and equipped with Explainable AI components. The model is evaluated using MAE, MSE, RMSE, R-Squared, Sensitivity, Specificity, and F1-Score as performance metrics. The results show that the Enhanced STING method achieved a MAE of 0.090 and RMSE of 0.0109, outperforming existing methods such as KNN, MICE, and Mean Imputation. Meanwhile, the Enhanced Kernel method achieved an F1-Score of 0.825. The SKDL method achieved MAE of 0.0870, MSE of 0.0175, RMSE of 0.0040, and R-Squared of 0.3367, which shows the best accuracy results compared to the Advanced Stage STING and Advanced Stage Kernel Methods. These findings indicate that the proposed approach not only produces more accurate imputations but also provides clinically interpretable outputs in the form of categorical data representing patient condition status. This research provides a significant contribution to the field of medical data science, particularly in the development of AI-based decision support systems for intensive care units.

Keywords: Imputation, Time Series, Multivariate, Data, STING, Kernel, Deep Level, Explainable

KATA PENGANTAR

Puji syukur dipanjatkan kehadirat Allah SWT, atas limpahan rahmat dan hidayahnya penulis dapat menyelesaikan Disertasi dengan judul “Imputasi Data Deret Waktu Multivariat Menggunakan Metode SKDL (STING Kernel Deep Level) With Explainable”. Penghargaan dan ucapan terima kasih penulis yang tulus dan mendalam kepada semua pihak yang telah berperan dan mendukung dalam proses penyelesaian Studi Doktor. Dalam kesempatan ini penulis ucapkan terima kasih yang sedalam-dalamnya kepada yth:

1. Prof. Dr. Ir. Siti Nurmaini, MT, selaku Promotor
2. Prof. Dr. Erwin, S.Si., M.Si, selaku Ko-Promotor
3. Prof. Dr. Ir. H. Nukman, M.T, selaku Koordinator Program Studi Doktor Ilmu-Ilmu Teknik Fakultas Teknik Program Pascasarjana Universitas Sriwijaya
4. Dr. Ir. Bhakti Yudho Suprpto, S.T., M.T., IPM selaku Dekan Fakultas Teknik Universitas Sriwijaya
5. Yuni Erika, selaku Administrasi Program Studi Doktor Ilmu Teknik

Secara khusus penulis menyampaikan terima kasih kepada Prof. Dr. Ir. Siti Nurmaini, MT atas dorongan dan dukungan moral dan intelektual yang diberikan selama ini dan khususnya dukungan untuk merampungkan Disertasi ini. Juga kepada Prof. Dr Erwin, S.Si.,M.Si yang terus memberikan dorongan dan masukan berharga untuk kesempurnaan Disertasi ini. Ucapan rasa terima kasih juga penulis ucapkan kepada Universitas Dehasen Bengkulu yang telah memberikan izin untuk melanjutkan studi Doktor. Terima kasih juga kepada rekan-rekan dan mahasiswa di Jurusan Informatika FILKOM Unived atas dukungan yang diberikan.

Akhirnya penghargaan dan terima kasih dengan penuh tulus penulis ucapkan kepada kedua orang tua (Aguswan dan Darmilawati), Istri (Mira Apriani), dan keluarga besar, yang tiada henti memberikan dorongan, dukungan, dan motivasi. Disertasi ini didedikasikan kepada Muhammad Tanzul dan Mira Tasya Rahma Dani, dengan harapan dapat menjadi inspirasi dan dorongan bagi mereka untuk meraih cita-cita di masa yang akan datang. Penulis mengakui masih banyak kekurangan dan ketidaksempurnaan yang ditemukan dalam Disertasi ini, yang sepenuhnya menjadi tanggung jawab penulis secara pribadi.

Salam Hormat,

Arius Satoni K

HALAMAN PERSEMBAHAN

Karya sederhana ini kupersembahkan...

**Kepada Allah SWT, Sang Maha Cahaya,
yang menuntunku dalam gelap, menguatkanmu dalam lelah,
dan menyelimutiku dengan rahmat-Nya yang tak bertepi.**

**Kepada Ayah dan Ibu tercinta,
yang setiap tetes doa dan pengorbanannya adalah jembatan yang
mengantarkanku hingga ke titik ini.
Tiada kata yang mampu membalas, kecuali doa agar kelak pahala ini kembali
pada kalian dalam bentuk yang paling indah.**

**Kepada istriku tersayang,
yang kesabaran dan cintanya menjelma menjadi pelabuhan hati,
tempatku bersandar ketika badai datang,
dan sumber kekuatan yang membuatku mampu bertahan.**

**Kepada keluargaku tercinta,
yang selalu menjadi rumah bagi lelahku,
dan alasan terindah untuk terus menatap masa depan dengan harapan.**

**Kepada para dosen, pembimbing, dan sahabat seperjuangan,
yang dengan keikhlasan dan kebijaksanaannya telah menuntunku,
menjadi cahaya kecil yang menerangi jalan panjang penuh kerikil ini.**

**Semoga setiap kata dalam Disertasi ini bukan hanya menjadi ilmu,
tetapi juga doa, amal, dan persembahan cinta yang akan terus hidup selamanya.**

DAFTAR ISI

ENDORSEMENT PAGE	i
HALAMAN PERSETUJUAN	ii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
HALAMAN PERSEMBAHAN	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
DAFTAR LAMPIRAN	xii
BAB 1 PENDAHULUAN	1
Latar Belakang Masalah	1
Rumusan Masalah	4
Tujuan Penelitian	5
Batasan Masalah	6
Sistematika Disertasi	6
BAB 2 TINJAUAN PUSTAKA	8
Artificial Intelligence	8
Machine Learning	11
Deep Learning	14
Imputasi	16
Algoritma Deep Learning dalam Imputasi	19
Penelitian Sejenis yang terkait dengan Metode Deep Learning	23
Arsitektur Auto Encoder	28
Metode STING	32
Metode Kernel	37
Data Vital Sign Pasien pada Unit Perawatan Intensif	42
Data Deret Waktu Multivariat	45
Metrik Evaluasi Model	48
BAB 3 METODOLOGI PENELITIAN	51
Kerangka Penelitian	51
Tahap Data Acquisition	52
Pre-Processing Data	61
Menghilangkan Data pada Data yang Nilainya hilang	65
Normalisasi Data	66
Pembagian/Splitting Data	68
Segmentasi Data	71
Pembuatan Nilai Hilang Buatan (Create Artificial NaN)	72
Data Label	74
Data NaN	76

Data Mask.....	78
Imputattion With Mean	80
Model atau Desain Imputasi Data Time Series Multivariat dengan Metode STING dan Kernel Tahap Lanjutan dengan Penjelasan.....	82
Hasil dari Model sebuah Metode STING dan Kernel Tahap Lanjutan dengan Penjelasan.....	87
Evaluasi Kinerja Model	88
BAB 4 HASIL DAN PEMBAHASAN	94
Hasil Penelitian	94
Pembahasan.....	104
Evaluasi Kinerja Model	114
BAB 5 PENUTUP	124
Kesimpulan	124
Saran.....	124
DAFTAR PUSTAKA	126

DAFTAR GAMBAR

Gambar 2.1.	Arsitektur Imputasi dengan Metode STING.....	34
Gambar 3.1.	Kerangka Kerja Pembuatan Program Imputasi Data Time Series Multivariat dengan menggunakan Metode STING dan Kernel Tahap Lanjutan serta Metode STING Kernel Deep Level (SKDL)	52
Gambar 3.2	Pre-Processing Data	62
Gambar 3.3	Contoh data label	75
Gambar 3.4	Contoh posisi nilai NaN pada data NaN	77
Gambar 3.5	Nilai yang hilang diisi dengan metode rata-rata	81
Gambar 3.6	Model atau Desain Metode Imputasi Data Time Series Multivariat (Metode STING dan Kernel Tahap Lanjutan serta Metode SKDL dengan Penjelasan)	82
Gambar 3.7	Arsitektur Metode STING Tahap Lanjutan	84
Gambar 3.8	Flowchart Implementasi Metode Kernel Tahap Lanjutan	86
Gambar 3.9	Hasil dari Model Sebuah Metode STING dan Kernel Tahap Lanjutan serta Metode SKDL dengan Penjelasan	88
Gambar 4.1.	Visualisasi Hasil Imputasi Data Mimic IV menggunakan Metode Kernel Tahap Lanjutan	102
Gambar 4.2.	Tren Kinerja Metode STING dan STING Tahap Lanjutan berdasarkan Nilai Metrik MAE	108
Gambar 4.3.	Perbandingan Performa Metode Kernel dan Kernel Tahap Lanjutan	112
Gambar 4.4.	Hasil Evaluasi Metode STING Tahap Lanjutan Menggunakan Matriks Evaluasi MAE, MSE, RMSE dan R-Squared	116
Gambar 4.5.	Perbandingan Nilai MAE dan RMSE pada Lima Metode Imputasi.....	121

DAFTAR TABEL

Tabel 2.1	Penelitian Sejenis yang terkait dengan Metode Deep Learning dalam Imputasi	24
Tabel 3.1	Tabel Database MIMIC-IV	54
Tabel 3.2	Tabel Chartevents MIMIC-IV	56
Tabel 3.3	Total data pada setiap kolom tanda vital pasien ICU.....	57
Tabel 3.4	Sampel data tanda vital pasien	57
Tabel. 3.5	Persentase Atribut Data Hilang pada Data Vital Sign Mimic-IV	59
Tabel 3.6	Contoh Sampel data setelah nilai NaN dibersihkan.....	66
Tabel 3.7	Contoh hasil Normalisasi Data	68
Tabel 3.8	Pembagian Data train dan Data valid pada Data Vital Sign Mimic-IV	69
Tabel 3.9	Perbandingan MAE, MSE, RMSE, R^2	92
Tabel 4.1.	Data Matrix Mimic IV	94
Tabel 4.2.	Mask Matrix data Mimic IV	95
Tabel 4.3.	Random Matrix data Mimic IV	95
Tabel 4.4.	Data Awal Mimic IV sebelum dilakukan Imputasi	96
Tabel 4.5.	Hasil Imputasi Data Mimic IV menggunakan Metode STING Tahap Lanjutan	97
Tabel 4.6.	Tampilan Hasil Imputasi Data Mimic IV Menggunakan Metode STING Tahap Lanjutan pada Aplikasi Phytion	98
Tabel 4.7	Hasil imputasi dari Metode STING Tahap Lanjutan Dalam bentuk data Kategorikal	99
Tabel 4.8.	Hasil Imputasi Data Mimic IV menggunakan Metode Kernel Tahap Lanjutan	101
Tabel 4.9.	Hasil Imputasi Data Mimic IV menggunakan Metode SKDL with Explainable	103
Tabel. 4.10.	Perbandingan Hasil Evaluasi Matrix dari Metode STING dan STING Tahap Lanjutan.....	106
Tabel. 4.11.	Perbandingan Hasil Matrix Evaluasi dari Metode Kernel dan Kernel Tahap Lanjutan.....	110
Tabel 4.12.	Hasil Evaluasi Metode STING Tahap Lanjutan Menggunakan Matrik Evaluasi MAE, MSE, RMSE dan R-Squared	115
Tabel. 4.13.	Hasil Matrix Evaluasi dari Metode Kernel Tahap Lanjutan.....	117
Tabel 4.14.	Analisis Distribusi Statistik Deskriptif Data Vital Sign Per Variabel.....	117
Tabel 4.15.	Perbandingan Kinerja Model Imputasi pada Evaluasi Komparatif dengan Metode Existing	119
Tabel 4.16.	Hasil Evaluasi Metrik Metode SKDL With Explainable.....	122
Tabel 4.17.	Perbandingan Hasil Evaluasi Metrik Metode SKDL With Explainable, Metode STING Tahap Lanjutan dan Metode Kernel Tahap Lanjutan	123

DAFTAR LAMPIRAN

- 1. Data Set**
- 2. Phyton Source Code**
- 3. Paper Jurnal Penelitian Q3**

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Data deret waktu multivariat adalah jenis data yang sangat penting dan ada diberbagai bidang kehidupan, di bidang finansial, kesehatan, lalu lintas [1], otomotif, domain jaringan [2], sains, teknik, dan aplikasi industri [3]. Data Deret waktu Multivariat banyak digunakan dalam pemantauan dan pemeliharaan industri, pemantauan Kesehatan, prakiraan cuaca, dan bidang lainnya [4]. Namun, tidak dapat dihindari bahwa data deret waktu bisa saja terdapat nilai yang hilang karena beberapa alasan, misalnya data yang dikumpulkan atau catatan yang dikumpulkan hilang karena kerusakan peralatan atau kesalahan komunikasi [5]. Data yang hilang dapat menghilangkan informasi penting dari atribut, namun tidak dapat langsung dihapus pada dataset. Untuk menangani data yang hilang terdapat beberapa cara antara lain penghapusan, imputasi dengan mean, mode, atau dengan metode prediksi [84].

Berurusan dengan nilai-nilai yang hilang dan deret waktu yang tidak lengkap adalah tantangan yang umum ditemui dalam aplikasi dunia nyata. Untuk memprediksi nilai yang hilang diperlukan sebuah pendekatan yang jelas, yaitu metode imputasi. Baru-baru ini, sejumlah besar metode imputasi bermunculan. Contohnya, metode BRITS yang dibangun berdasarkan modul berulang dua arah, serta *Generative Adversarial Imputation Networks (GAIN)*, salah satu pendekatan awal yang menggunakan pelatihan adversarial untuk tugas imputasi [6]. Penerapan Imputasi data deret waktu multivariat sudah banyak diimplementasikan di berbagai bidang, contohnya pada aplikasi ilmu lingkungan, sebagai contohnya adalah kondisi aliran tinggi atau fluktuasi dapat mempengaruhi siklus karbon dan nukrin dilingkungan zona hiporeik [7]. Di bidang Kesehatan, contoh penerapan imputasi untuk memprediksi hasil klinis seperti kematian, dekompensasi, lama rawat inap, dan risiko penyakit, dari data rangkaian waktu multivariat yang kompleks dapat memfasilitasi pengelolaan unit perawatan kritis yang efektif dan

rekomendasi pengobatan otomatis yang dipersonalisasi untuk pasien [8].

Banyak penelitian tentang imputasi nilai yang hilang hanya berfokus pada data tabular multivariat tanpa mempertimbangkan variabel yang berubah terhadap waktu [9]. Padahal, banyak dataset nyata dalam ilmu kesehatan dan aplikasi sensor elektronik mencakup data sekuensial dengan variasi waktu [9]. Dalam catatan kesehatan elektronik, data pasien dikumpulkan dari waktu ke waktu. Hal ini mengakibatkan munculnya nilai yang hilang, baik pada aspek variabel (*cross-sectional*) maupun waktu (*longitudinal*). Dengan demikian, diperlukan metode imputasi yang mampu mengisi data hilang secara akurat [9]. Metode berbasis *Recurrent Neural Network* (RNN) telah menjadi solusi utama untuk menangani data deret waktu klinis, karena RNN dapat mengolah data sekuensial dengan panjang bervariasi. Namun, metode RNN konvensional biasanya dirancang untuk menangani data dengan interval waktu yang konstan, sehingga kinerjanya kurang optimal ketika menghadapi interval waktu yang tidak teratur [10]. Salah satu metode imputasi terbaru adalah *Self-Attention based Time Series Imputation Network using GAN* (STING), yang memanfaatkan jaringan adversarial generatif dan jaringan saraf berulang dua arah untuk mempelajari representasi laten deret waktu. Namun, metode STING memiliki keterbatasan karena tidak dapat menangani jenis data lain, seperti data kategorikal atau kualitatif [1]. Selain metode STING, terdapat juga metode kernel, yaitu pendekatan dengan strategi pembelajaran ansambel berbasis campuran Bayesian yang dapat mengeksploitasi informasi yang hilang tanpa harus melalui proses imputasi tradisional. Akan tetapi, kelemahan metode kernel adalah tidak mampu mengeksploitasi pola dari informasi yang hilang, padahal pola tersebut sering muncul dalam data medis [11]. Untuk mengatasi kelemahan masing-masing metode, diperlukan pengembangan lanjutan (tahap 2) yang disebut sebagai imputasi tahap mendalam (Deep Level).

Penelitian [11] menyimpulkan bahwa metode kernel sangat cocok digunakan pada data catatan kesehatan elektronik. Sementara itu, penelitian [1] menunjukkan bahwa metode STING mampu mengungguli pendekatan lain dalam hal akurasi imputasi serta performa pada tugas hilir. Oleh karena itu, penelitian ini

berfokus pada eksperimen dengan data kesehatan, khususnya data vital sign MIMIC-IV. Imputasi pada data yang hilang diharapkan dapat memudahkan analisis kondisi pasien dan mendukung keputusan dokter secara lebih akurat. Penelitian ini juga mengembangkan metode kernel untuk imputasi pada data vital sign MIMIC-IV secara lebih mendalam sehingga diharapkan menghasilkan akurasi yang lebih tinggi dibandingkan metode kernel sebelumnya. Riset [12] menjelaskan bahwa kernel cluster deret waktu menyediakan alat yang ampuh untuk menganalisis deret waktu multivariat dengan data hilang. Sementara metode STING terbukti efektif untuk imputasi pada data tabular di bidang kesehatan, akan tetapi kelemahan metode STING selama ini adalah tidak dapat menghasilkan imputasi dalam bentuk data kategorikal [1].

Untuk mengatasi kelemahan tersebut dan dapat diterapkan pada data vital sign MIMIC-IV, diperlukan metode imputasi baru yang lebih akurat serta mampu menghasilkan data kategorikal. Oleh karena itu, penulis mengusulkan metode gabungan STING dan Kernel Tahap Lanjutan dengan Explainability. Dengan metode ini diharapkan dapat dihasilkan imputasi yang lebih mendalam, akurat, serta memberikan penjelasan (explainable) atas hasil imputasi. Penjelasan ini penting untuk menunjukkan bahwa hasil imputasi dapat dipertanggungjawabkan, sejalan dengan riset [13] yang menekankan pentingnya model yang dapat diinterpretasikan. Selain itu, riset [14] menunjukkan pentingnya kemampuan melakukan imputasi pada data lintas dimensi dengan penerapan algoritma Clinical Decision Support Algorithm (CDSA) yang dapat melibatkan berbagai modalitas data. Penelitian [15] juga menjelaskan tentang imputasi berbasis optimasi dengan menghubungkan pola nilai hilang pada data yang saling terkait. Oleh karena itu, penelitian ini mengembangkan metode baru, yakni metode STING Kernel Deep Level With Explainable/SKDL With Explainable (Metode STING dan Kernel Tahap Lanjutan serta Metode SKDL dengan Penjelasan), untuk imputasi data time series Multivariat di bidang kesehatan.

Berdasarkan uraian tersebut, jelas bahwa tantangan utama dalam pengolahan data vital sign multivariat, khususnya pada dataset MIMIC-IV, terletak pada kompleksitas pola nilai hilang, kebutuhan akan akurasi imputasi

yang tinggi, serta pentingnya keterjelasan (explainability) hasil prediksi. Penggunaan metode konvensional terbukti belum mampu memberikan hasil yang optimal, sementara model deep learning yang ada sering kali dipandang sebagai black box sehingga sulit dipertanggungjawabkan dalam konteks medis.

Oleh karena itu, pengembangan metode STING Kernel Deep Level With Explainable (SKDL With Explainable) menjadi relevan dan signifikan. Metode ini diharapkan tidak hanya mampu meningkatkan akurasi imputasi, tetapi juga memberikan penjelasan yang transparan, sehingga hasil yang diperoleh dapat dipercaya dan digunakan sebagai dasar pengambilan keputusan klinis. Dengan demikian, penelitian ini memiliki urgensi akademik dan praktis yang kuat untuk berkontribusi pada kemajuan ilmu pengetahuan serta mendukung implementasi teknologi kecerdasan buatan yang etis dan dapat diandalkan dalam bidang kesehatan.

1.2 Rumusan Masalah

Beberapa metode Imputasi sudah dikembangkan untuk mengisi nilai yang hilang pada data kesehatan, akan tetapi masih memiliki kelemahan dan keterbatasan. Beberapa penelitian sudah mengembangkan metode STING dan Kernel untuk sebuah kasus imputasi, yang masih dapat dikembangkan untuk menghasilkan sebuah imputasi yang lebih akurat dan sesuai dengan kebutuhan. Untuk itu pertanyaan riset yang akan dijawab pada penelitian ini adalah:

“Bagaimana menghasilkan sebuah metode imputasi yang baru yakni, metode STING dan Kernel Tahap Lanjutan serta Metode SKDL dengan Penjelasan”

Untuk menyelesaikan permasalahan tersebut, beberapa issue penting yang harus dipertimbangkan antara lain:

1. Bagaimana merancang metode STING Tahap Lanjutan yang dapat menghasilkan imputasi yang lebih baik dari Metode STING sebelumnya yang tidak dapat menghasilkan data katagorikal?

2. Bagaimana merancang metode Kernel Tahap Lanjutan yang dapat menghasilkan nilai akurasi yang lebih baik dibandingkan dengan metode kernel pada riset sebelumnya?
3. Bagaimana merancang metode STING dan Kernel Tahap Lanjutan dengan Penjelasan yang dapat menghasilkan imputasi pada dua data dengan dimensi yang berbeda dan dapat juga menghasilkan sebuah penjelasan?
4. Bagaimana memvalidasi dan mengevaluasi metode STING dan Kernel Tahap Lanjutan dengan Penjelasan, sehingga menghasilkan sebuah teori Imputasi yang lebih baik dan relevan dengan kebutuhan saat ini.
5. Bagaimana merancang metode SKDL with Explainable yang dapat menghasilkan nilai akurasi yang lebih baik dibandingkan dengan metode STING Tahap Lanjutan dan Kernel Tahap Lanjutan?

1.3 Tujuan Penelitian

Secara umum, penelitian ini akan memberikan model pendekatan yang dapat diterapkan di instansi kesehatan sehingga imputasi baru yang dirancang dapat membantu pihak instansi untuk mendapatkan hasil analisis yang terbaik dan juga keputusan dokter yang tepat untuk diberikan kepada pasien. Kontribusi penelitian ini akan memberikan model imputasi yang terbaik dan sesuai dengan kebutuhan pihak instansi kesehatan.

Untuk mencapai hal tersebut, maka tujuan dari penelitian ini adalah:

1. Merancang metode STING Tahap Lanjutan yang dapat menghasilkan imputasi yang lebih baik dibandingkan dengan Metode STING sebelumnya.
2. Merancang metode Kernel Tahap Lanjutan yang dapat menghasilkan nilai akurasi yang lebih baik dibandingkan dengan metode Kernel pada riset sebelumnya.
3. Merancang metode STING dan Kernel Tahap Lanjutan dengan Penjelasan yang dapat menghasilkan imputasi pada dua data dengan dimensi yang berbeda dan dapat juga menghasilkan sebuah penjelasan.

4. Merancang metode STING dan Kernel Tahap Lanjutan dengan Penjelasan, sehingga menghasilkan sebuah teori Imputasi yang lebih baik dan relevan dengan kebutuhan saat ini.
5. Merancang metode SKDL with Explainable yang dapat menghasilkan nilai akurasi yang lebih baik dibandingkan dengan metode STING Tahap Lanjutan dan Kernel Tahap Lanjutan.

1.4 Batasan Masalah

Berikut ini beberapa batasan penelitian yang ditentukan:

1. Data penelitian hanya data vital sign Mimic-IV.
2. Hanya membahas 3 metode imputasi, yakni Metode STING Tahap Lanjutan, Kernel Tahap Lanjutan dan SKDL (STING Kernel Deep Level).
3. Pengujian metode imputasi yang dibangun dengan menggunakan metric evaluasi Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE) dan R-squared.
4. Penjelasan pada model hanya menjelaskan terhadap hasil Imputasi yang didapatkan.
5. Hanya membahas tentang Data Time Series.

1.5 Sistematika Disertasi

Adapun Sistematika Penelitian Disertasi ini, yaitu:

BAB I PENDAHULUAN

Dalam Bab I terdiri dari Latar Belakang Penelitian, Perumusan Masalah, Tujuan Penelitian, ruang lingkup penelitian dan Sistematika Penulisan Disertasi.

BAB II TINJAUAN PUSTAKA

Dalam bab ini berisikan Teori mendasar tentang Imputasi, Imputasi Multivariat, Deep Learning dalam Imputasi, Auto Encoder, Metode STING,

Metode Kernel, Evaluasai Model dan Penelitian Sejenis.

BAB III METODOLOGI PENELITIAN

Dalam bab ini berisikan tentang langkah kerja Penelitian, yang dimulai dari Studi Literatur, Pra-Pemrosesan Dataset, Membangun Metode/Model Imputasi terbaru, Evaluasi dan Pengujian model.

BAB IV HASIL DAN PEMBAHASAN

Dalam bab ini berisikan tentang hasil Penelitian, Pembahasan dan Pengujian Model dari Penelitian ini.

BAB V PENUTUP

Dalam bab ini berisikan tentang Kesimpulan dan Saran dari Penelitian serta merupakan bagian penutup dari Disertasi ini.

DAFTAR PUSTAKA

- [1] E. Oh, T. Kim, Y. Ji, and S. Khyalia, "STING : Self-attention based Time-series Imputation Networks using GAN."
- [2] B. P. Singh *et al.*, "Multi-resolution Networks For Flexible Irregular Time Series Modeling (Multi-FIT)."
- [3] G. et al Zerveas, "A TRANSFORMER-BASED FRAMEWORK FOR MULTI-VARIATE TIME SERIES REPRESENTATION LEARNING," pp. 1–20, 2020.
- [4] R. Wang, Z. Zhang, Q. Wang, and J. Sun, "TLGRU : time and location gated recurrent unit for multivariate time series imputation," *EURASIP J. Adv. Signal Process.*, 2022.
- [5] I. M. By and G. R. N. Eural, "FILLING THE GAPS : MULTIVARIATE TIME SERIES," pp. 1–20, 2022.
- [6] D. Wang, Y. Yan, R. Qiu, Y. Zhu, and H. Tong, "Networked Time Series Imputation via Position-aware Graph Enhanced Variational Autoencoders."
- [7] J. Park *et al.*, "Imputasi nilai hilang jangka panjang untuk data deret waktu menggunakan jaringan saraf dalam," vol. 6, pp. 9071–9091, 2023.
- [8] S. Tipirneni and V. Tech, "Self-Supervised Transformer for Sparse and Irregularly Sampled Multivariate Clinical Time-Series," *ACM Trans. Knowl. Discov. Data*, vol. 1, no. 1, pp. 1–18, 2022.
- [9] M. Kazijevs and M. D. Samad, "Deep Imputation of Missing Values in Time Series Health Data : A Review with Benchmarking," no. Mvi, 2023.
- [10] Y. Lee, E. Jun, and H. Suk, "Multi-view Integration Learning for Irregularly-sampled Clinical Time Series," 2020.
- [11] C. Soguero-ruiz and R. Jenssen, "A Kernel to Exploit Informative Missingness in Multivariate Time Series from EHRs," no. 9037, 2019.
- [12] K. Øyvind, C. Soguero-ruiz, F. Maria, A. Revhaug, and R. Jenssen, "Time series cluster kernels to exploit informative missingness and incomplete label information," vol. 115, 2021.
- [13] W. Jung, E. Jun, H. Suk, and D. Neuroimaging, "NeuroImage Deep recurrent model for individualized prediction of Alzheimer ' s disease progression," *Neuroimage*, vol. 237, no. March, p. 118143, 2021.
- [14] J. Ma and Z. Shou, "CDSA: Cross-Dimensional Self-Attention for Multivariate, Geo-tagged Time Series Imputation," pp. 1–11, 2018.
- [15] D. Aureli, R. Bruni, and C. Daraio, "MethodsX Optimization methods for the imputation of missing values in Educational Institutions Data," *MethodsX*, vol. 8, p. 101208, 2021.
- [16] R. K. dan M. C. Mateusz Szczepański, Marek Pawlicki, "Metode Tingkat Lanjut untuk Menangani Nilai yang Hilang di Deteksi Intrusi Jaringan," vol. 0, pp. 1–23, 2022.
- [17] N. H. Abidatul Izzah, "IMPUTASI MISSING DATA MENGGUNAKAN METODE K-NEAREST NEIGHBOUR DENGAN OPTIMASI ALGORITMA GENETIKA."
- [18] Q. Ni and X. Cao, "MBGAN : Jaringan permusuhan generatif yang ditingkatkan dengan perhatian mandiri multi-kepala dan RNN dua arah untuk imputasi deret waktu," vol. 115, no. April, 2022.
- [19] Y. dkk Zhang, "Imputasi nilai tidak ada dalam rangkaian waktu multivariat dengan jaringan permusuhan generatif ujung ke ujung," vol. 551, pp. 67–82, 2021.
- [20] J. Zhang, "Imputasi Data Hilang Rangkaian Waktu Multivariat Menggunakan Autoencoder Denoising Berulang," no. November, 2019.

- [21] J. Q. Zhao, C. Rong, and C. Bin Lin, "Imputasi data deret waktu multivariat menggunakan mekanisme berbasis perhatian," vol. 542, 2023.
- [22] Liu, Yan Du, Wenjie Côté, David, "SAITS: Imputasi berbasis perhatian diri untuk deret waktu," vol. 219, 2023.
- [23] X. Zhang, C. Yan, C. Gao, B. A. Malin, and Y. Chen, "Predicting Missing Values in Medical Data Via XGBoost Regression," *J. Healthc. Informatics Res.*, vol. 4, no. 4, pp. 383–394, 2020, doi: 10.1007/s41666-020-00077-1.
- [24] Y. P. Chen, C. H. Huang, Y. H. Lo, Y. Y. Chen, and F. Lai, "Combining attention with spectrum to handle missing values on time series data without imputation," *Inf. Sci. (Ny)*, vol. 609, no. August, pp. 1271–1287, 2022, doi: 10.1016/j.ins.2022.07.124.
- [25] R. L. Hale, "Cluster analysis in school psychology: An example," *J. Sch. Psychol.*, vol. 19, no. 1, pp. 51–56, 1981, doi: 10.1016/0022-4405(81)90007-8.
- [26] Z. Che, S. Purushotham, K. Cho, D. Sontag, and Y. Liu, "Recurrent Neural Networks for Multivariate Time Series with Missing Values," *Sci. Rep.*, vol. 8, no. 1, pp. 1–13, 2018, doi: 10.1038/s41598-018-24271-9.
- [27] X. Wang and C. Wang, "Time Series Data Cleaning: A Survey," *IEEE Access*, vol. 8, pp. 1866–1881, 2020, doi: 10.1109/ACCESS.2019.2962152.
- [28] M. Chen, H. Zhu, Y. Chen, and Y. Wang, "A Novel Missing Data Imputation Approach for Time Series Air Quality Data Based on Logistic Regression," 2022.
- [29] D. Chang, D. Chang, and M. Pourhomayoun, "Risk prediction of critical vital signs for ICU patients using recurrent neural network," *Proc. - 6th Annu. Conf. Comput. Sci. Comput. Intell. CSCI 2019*, pp. 1003–1006, 2019, doi: 10.1109/CSCI49370.2019.00191.
- [30] A. Youssef *et al.*, "Vital Signs Prediction and Early Warning Score Calculation Based on Continuous Monitoring of Hospitalised Patients Using Wearable Technology," 2020, doi: 10.3390/s20226593.
- [31] A. L. Goldberger *et al.*, "Current Perspective," 2000.
- [32] B. Shickel, P. J. Tighe, A. Bihorac, and P. Rashidi, "Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis," *IEEE J. Biomed. Heal. Informatics*, vol. 22, no. 5, pp. 1589–1604, 2018, doi: 10.1109/JBHI.2017.2767063.
- [33] Johnson, A.E.W., Bulgarelli, L., Shen, L. et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data* 10, 1 (2023). <https://doi.org/10.1038/s41597-022-01899-x>
- [34] Aiman Nadir R dkk, Sistem Prediksi Harga Emas Berdasarkan Data *Time Series* Menggunakan Metode *Artificial Neural Network* (ANN), <https://doi.org/10.47709/digitech.v3i2.2877>
- [35] D. E. J. H. H. Goldberg, "Machine Learning. Vol. 3, Nos. 2-3, Special Issue on Genetic Algorithms.," 1988.
- [36] G.-B. Q.-Y. Z. C.-K. S. Huang, "Extreme learning machine: theory and applications.," *Neurocomputing*, pp. 489-501, 2006.
- [37] M. C. P. T. S. S. V. Somvanshi, "A review of machine learning techniques using decision tree and support vector machine.," in international conference on computing communication control and automation (ICCUBE) (pp. 1-7). IEEE., 2016.
- [38] A. Annunz, "Analisis Runtun Waktu Untuk Memprediksi Jumlah Mahasiswa Baru Dengan Model Prophet Facebook," *semanticscholar.org*, 2021.
- [39] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *nature*, 521(7553), 436-444.
- [40] Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). Deep face recognition.
- [41] Berniker, M., & Kording, K. P. (2015). Deep networks for motor control functions.

Frontiers in computational neuroscience, 9, 32.

- [42] Bae, H. S., Lee, H. J., & Lee, S. G. (2016, June). Voice recognition based on adaptive MFCC and Deep Learning. In 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA) (pp. 1542-1546). IEEE.
- [43] Kruger, N., Janssen, P., Kalkan, S., Lappe, M., Leonardis, A., Piater, J., ... & Wiskott, L. (2012). Deep hierarchies in the primate visual cortex: What can we learn for computer vision?. *IEEE transactions on pattern analysis and machine intelligence*, 35(8), 1847-1871.
- [44] P. B. Weerakody, K. W. Wong, G. Wang, and W. Ela, "A review of irregular time series data handling with gated recurrent neural networks," *Neurocomputing*, vol. 441, pp. 161–178, 2021, doi: 10.1016/j.neucom.2021.02.046.
- [45] J. Yoon, W. R. Zame, and M. Van Der Schaar, "Multi-directional Recurrent Neural Networks : A Novel Method for Estimating Missing Data," *ICML Work.*, 2017.
- [46] C. Fang and C. Wang, "Time Series Data Imputation: A Survey on Deep Learning Approaches," 2020, [Online]. Available: <http://arxiv.org/abs/2011.11347>.
- [47] H. Khan, X. Wang, and H. Liu, "Handling missing data through deep convolutional neural network," *Inf. Sci. (Ny).*, vol. 595, pp. 278–293, 2022, doi: 10.1016/j.ins.2022.02.051.
- [48] Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons & Fractals*, 135, 109864. <https://doi.org/10.1016/j.chaos.2020.109864>
- [49] Nagarajan, G, dkk "Imputasi data yang hilang pada data biomedis menggunakan pengelompokan yang dipelajari secara mendalam dan regresi teregulasi L2 berdasarkan ketidakpastian simetris," *Kecerdasan Buatan dalam Kedokteran.*, Vol. 123 No. 102214, 2022
- [50] Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine Learning in Medicine. *New England Journal of Medicine*, 380(14), 1347–1358. <https://doi.org/10.1056/nejmra1814259>
- [51] Masrichah, Siti. (2023). Khatulistiwa: Jurnal Pendidikan dan Sosial Humaniora. Vol.3, No.3, September 2023 e-ISSN: 2962-4010; pISSN: 2962-4444, Hal 83-101 DOI: <https://doi.org/10.55606/khatulistiwa.v3i3.1860>
- [52] Utami, 2021. <https://www.kompas.com/skola/read/2021/07/05/121323869/artificial-intelligence-ai-pengertian-perkembangan-cara-kerja-dan>
- [53] Pakpahan, Roida, 2021. Analisa Pengaruh Implementasi Artificial Intelligence Dalam Kehidupan Manusia. *Journal of Information System, Informatics and Computing*. e-ISSN : 2597-3673 (Online), p-ISSN: 2579-5201 (Printed) Vol.5 No.2 Desember 2021
- [54] INFORMATIKA, C., & Yudi Herdiana. (2022). PENERAPAN MACHINE LEARNING DENGAN MODEL LINEAR REGRESSION TERHADAP ANALISIS KUALITAS HASIL PETIK THE DI PT. PERKEBUNAN NUSANTARA VIII KEBUN SEDEP. *COMPUTING | Jurnal Informatika*, 9(01), 1–9. Retrieved from <https://ejournal.unibba.ac.id/index.php/computing/article/view/855>
- [55] Robby Takdirillah. 2020. Apa itu Machine Learning? Beserta Pengertian dan Cara Kerjanya. URL : <https://www.dicoding.com/blog/machine-learning-adalah/>. Diakses tanggal 7 Januari 2021
- [56] R. Abellera, and R. Bulusu, "Oracle Business Intelligence with Machine Learning: Artificial Intelligence Techniques in OBIEE for Actionable BI", NewYork: Apress, 2018
- [57] Fard, et al., "Application of Machine Learning in the Prediction of Covid-19 daily new cases: A Scoping Review.", *Heliyon*, 7(10), 2021.
- [58] Wardhana, Rakha Gusti dkk. PENERAPAN MACHINE LEARNING DALAM

PREDIKSI TINGKAT KASUS PENYAKIT DI INDONESIA. *Journal of Information System Management (JOISM)* e-ISSN: 2715-3088 Vol. 5, No. 1 (2023)

- [59] M.Z. Alom, T.M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M.S. Nasrin, M. Hasan, B.C. Van Essen, A.A.S. Awwal, dan V.K. Asari, "A State-of-the-Art Survey on Deep Learning Theory and Architectures," *Electronics*, Vol. 8, No. 3, hal. 1-3, 2019.
- [60] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview," *Neural Netw*, Vol. 61, hal. 85-117, 2015.
- [61] Y. LeCun, Y. Bengio, dan G. Hinton, "Deep Learning," *Nature*, Vol. 521, hal. 436-444, 2015.
- [62] Y. Bengio, A. Courville, dan P. Vincent, "Representation Learning: A Review and New Perspectives," *IEEE Trans. on Pattern Analysis and Machine Intel.*, Vol. 35, No. 8, hal. 1798-1828, 2013.
- [63] Diponegoro, M. Haris dkk. Tinjauan Pustaka Sistematis: Implementasi Metode *Deep Learning* pada Prediksi Kinerja Murid. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*. Vol. 10, No. 2, Mei 2021
- [64] S. Nikfalazar, C. H. Yeh, S. Bedingfield, and H. A. Khorshidi, "Missing data imputation using decision trees and fuzzy clustering with iterative learning," *Knowl Inf Syst*, vol. 62, no. 6, pp. 2419–2437, Jun. 2020, doi: 10.1007/s10115-019-01427-1.
- [65] M. Alabadla *et al.*, "Systematic Review of Using Machine Learning in Imputing Missing Values," *IEEE Access*, vol. 10. Institute of Electrical and Electronics Engineers Inc., pp. 44483–44502, 2022. doi: 10.1109/ACCESS.2022.3160841.
- [66] I. Pratama, A. E. Permanasari, I. Ardiyanto, and R. Indrayani, "A review of missing values handling methods on time-series data," in *2016 International Conference on Information Technology Systems and Innovation, ICITSI 2016 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., Feb. 2017. doi: 10.1109/ICITSI.2016.7858189.
- [67] N. Fazakis, G. Kostopoulos, S. Kotsiantis, and I. Mporas, "Iterative Robust Semi-Supervised Missing Data Imputation," *IEEE Access*, vol. 8, pp. 90555–90569, 2020, doi: 10.1109/ACCESS.2020.2994033.
- [68] R. Fajar, R. Setya Perdana, and Indriati, "Implementasi Metode Naïve Bayes Dengan Perbaikan Missing Value Menggunakan Metode Nearest Neighbor Imputation Studi Kasus: Penyakit Malaria Di Kabupaten Malang," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 8, pp. 2430–2434, Aug. 2018, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [69] Y. Kristian, I. K. E. Purnama, E. H. Sutanto, L. Zaman, E. I. Setiawan, and M. H. Purnomo, "Klasifikasi Nyeri pada Video Ekspresi Wajah Bayi Menggunakan DCNN Autoencoder dan LSTM," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 7, no. 3, pp. 308–316, 2018.
- [70] J. W. G. Putra, "Pengenalan Konsep Pembelajaran Mesin dan Deep Learning," pp. 1–235, 2019.
- [71] M. R. Fauzi *et al.*, "MODEL PREDIKSI CANCER MENGGUNAKAN AUTOENCODER," vol. 3, no. 1, 2019.
- [72] S. Prodi, T. Informatika, F. Informatika, and U. Telkom, "DENOISING SINYAL EKG MENGGUNAKAN DEEP NEURAL NETWORK DENGAN STACKED DENOISING AUTOENCODERS ECG SIGNAL DENOISING USING DEEP NEURAL NETWORK WITH STACKED DENOISING AUTOENCODERS," vol. 4, no. 3, pp. 5024–5030, 2017.
- [73] AYU AULIA, Prof. Subanar, Ph.D. REGRESI POLINOMIAL LOKAL DENGAN FUNGSI KERNEL GAUSSIAN Universitas Gadjah Mada, 2016. Diunduh dari <http://etd.repository.ugm.ac.id/>

- [74] Ahrens, T. 2008. The most important vital signs are not being measured. *Aust. Crit. Care* 21 (1), 3–5. (<https://www.sciencedirect.com/science/article/abs/pii/S1036731407002044>, diakses pada 9 Agustus 2019).
- [75] Forkan, Abdur, RM., Khalil, Ibrahim. 2017. PEACE Home: Probabilistic estimation of abnormal clinical events using vital sign correlations for reliable home-based monitoring. *Pervasive and Mobile Computing*, Volume 38, Part 2. (<https://reader.elsevier.com/reader/sd/pii/S1574119216304540?token=88C50266851AB5AA4FC85D8A6907FBA9229659E706CA9B028EED0210A54349D4FCADA506C9F3A20E253107FE76E87186>, diakses pada 26 September 2019).
- [76] Smith GB, Prytherch DR, Schmidt P, Featherstone PI, Knight D, Clements G, et al. 2006. Hospital-wide physiological surveillance-a new approach to the early identification and management of the sick patient. *Resuscitation*. (<https://www.sciencedirect.com/science/article/abs/pii/S0300957206001286?via%3Dihub>, diakses pada 26 September 2019).
- [77] Noviestasari, E., Supartini, Y. 2015. *Keperawatan Dasar: Manual Keterampilan Klinis Edisi Pertama*. Singapore: Elseiver Pte Ltd.
- [78] Sarotama, Afrias dkk. 2019. Implementasi Peringatan Abnormalitas Tanda-Tanda Vital pada Telemedicine Website: jurnal.umj.ac.id/index.php/semnastek
- [79] A. Hicham, P. A. Sabri, and P. H. Tairi, “A Survey on Educational Data Mining [2014-2019] 1 st,” in *2020 International Conference on Intelligent Systems and Computer Vision (ISCV)*, Fez, Morocco: IEEE, 2020. doi: 10.1109/ISCV49265.2020.9204013.
- [80] H. Benhar, A. Idri, and J. L Fernández-Alemán, “Data preprocessing for heart disease classification: A systematic literature review.,” *Computer Methods and Programs in Biomedicine*, vol. 195. Elsevier Ireland Ltd, Oct. 01, 2020. doi: 10.1016/j.cmpb.2020.105635.
- [81] G. Gunadi, “Penerapan Algoritma K-MEANS CLUSTERING Untuk Menganalisa Transaksi Penjualan Jasa Cetak Pada Unit Print On Demand (POD) Percetakan Gramedia,” *Infotech: Journal of Technology Information*, vol. 8, no. 2, pp. 117–126, Nov. 2022, doi: 10.37365/jti.v8i2.148.
- [82] Purwati Neni, Hendra Kurniawan, and Sri Karnila, *Data Mining*, vol. 1. Zahira Media Publisher, 2021.
- [83] Prasetya, M. Riko Anshori dkk. Penanganan Imputasi Missing Values pada Data Time Series dengan Menggunakan Metode Data Mining. *Jurnal Informasi dan Teknologi*. Vol. 5. No. 2. Hal: 56-62. e-ISSN: 2714-9730. <https://jidt.org/jidt> 2023
- [84] Desiani, Anita dkk. Penanganan Data Hilang Menggunakan Kombinasi Teknik Penghapusan, Mean, Mode dan Imputasi Jaringan Syaraf Tiruan untuk Dataset Penyakit Jantung. *Sains dan Teknologi Indonesia*. e-ISSN:2580-4391 p-ISSN:2580-4405. Vol. 6, Nomor 4, Oktober 2021. <https://doi.org/10.26554/sti.2021.6.4.303-312>
- [85] Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2018). Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. *IEEE Journal of Biomedical and Health Informatics*, 22(5), 1589–1604. <https://doi.org/10.1109/JBHI.2017.2767063>
- [86] Caruana, R., Lou, Y., Gehrke, J., Koch, P., Sturm, M., & Elhadad, N. (2015). Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission. *KDD '15: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. <https://doi.org/10.1145/2783258.2788613>
- [87] Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2017). Deep learning for

- healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*, 19(6), 1236–1246. <https://doi.org/10.1093/bib/bbx044>
- [88] Che, Z., Purushotham, S., Cho, K., Sontag, D., & Liu, Y. (2018). Recurrent Neural Networks for Multivariate Time Series with Missing Values. *Scientific Reports*, 8, 6085. <https://doi.org/10.1038/s41598-018-24271-9>
 - [89] Tonekaboni, S., Joshi, S., McCradden, M. D., & Goldenberg, A. (2019). What Clinicians Want: Contextualizing Explainable Machine Learning for Clinical End Use. *Proceedings of Machine Learning Research*, 106:359–380.
 - [90] Yoon, J., Zame, W., & van der Schaar, M. (2018). Estimating Missing Data in Temporal Data Streams using Multi-Directional Recurrent Neural Networks. *IEEE Transactions on Biomedical Engineering*, 66(5), 1477–1490.
 - [91] Gondara, L., & Wang, K. (2018). MIDA: Multiple Imputation using Denoising Autoencoders. *arXiv preprint arXiv:1705.02737*.
 - [92] Kingma, D. P., & Welling, M. (2014). Auto-Encoding Variational Bayes. *arXiv preprint arXiv:1312.6114*.
 - [93] Zhou, L., Zhang, W., Wu, Y., & Xu, W. (2023). STING: Self-attention Based Time-series Imputation Networks Using GAN. *IEEE Transactions on Neural Networks and Learning Systems*. <https://doi.org/10.1109/TNNLS.2023.3245567>
 - [94] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is All You Need. *Advances in Neural Information Processing Systems (NeurIPS)*, 30.
 - [95] Shawe-Taylor, J., & Cristianini, N. (2004). *Kernel Methods for Pattern Analysis*. Cambridge University Press.
 - [96] Tsutsui, I., Otsuka, A., Tanaka, Y., & Ikegami, T. (2018). Time Series Imputation Using Similarity Kernel with Application to Sensor Data. *IEEE Sensors Journal*.
 - [97] Zhao, L. et al. (2017). *Vital Sign Monitoring and ICU Early Warning Systems*. Journal of Clinical Monitoring and Computing.
 - [98] Subbe, C.P., et al. (2001). *Validation of physiological scoring systems*. QJM: An International Journal of Medicine.
 - [99] Clifton, D.A., et al. (2012). *Predictive Monitoring of Mobile Patients by Combining Clinical Observations with Data from Wearable Sensors*. IEEE Journal of Biomedical and Health Informatics.
 - [100] Smith, G.B., et al. (2008). *Hospital-wide physiological surveillance—a new approach to the early identification and management of the sick patient*. Resuscitation, 77(2), 157-163.
 - [101] Odell, M. (2010). *Detection and management of the deteriorating ward patient: a critical review of the literature*. Nursing in Critical Care, 15(1), 1-9.
 - [102] Potter, P.A. & Perry, A.G. (2016). *Fundamentals of Nursing*. Elsevier Health Sciences.
 - [103] Perry, A.G., Potter, P.A., & Ostendorf, W.R. (2017). *Clinical Nursing Skills and Techniques*. Elsevier Mosby.
 - [104] Chatfield, C. (2004). *The Analysis of Time Series: An Introduction*. Chapman and Hall/CRC.
 - [105] Box, G.E.P., Jenkins, G.M., Reinsel, G.C. (2015). *Time Series Analysis: Forecasting and Control*. Wiley.
 - [106] Little, R.J.A., & Rubin, D.B. (2019). *Statistical Analysis with Missing Data*. Wiley.
 - [107] Cao, W., Wang, D., Li, J., Zhou, H., Li, L., & Li, Y. (2018). *BRITS: Bidirectional Recurrent Imputation for Time Series*. In Proceedings of NeurIPS.
 - [108] Zhao, L. et al. (2017). *Vital Sign Monitoring in ICUs Using Time Series Models*. IEEE Transactions on Biomedical Engineering.
 - [109] Yoon, J., Zame, W. R., & van der Schaar, M. (2018). *Estimating Missing Data in*

- Temporal Data Streams Using Multi-directional Recurrent Neural Networks*. IEEE Transactions on Biomedical Engineering.
- [110] Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.).
 - [111] Batista, G. E. A. P. A., & Monard, M. C. (2002). *A study of K-nearest neighbour as an imputation method*. In Proceedings of the 2002 International Conference on Hybrid Intelligent Systems.
 - [112] White, I. R., Royston, P., & Wood, A. M. (2011). *Multiple imputation using chained equations: Issues and guidance for practice*. *Statistics in Medicine*, 30(4), 377–399.
 - [113] Che, Z., Purushotham, S., Cho, K., Sontag, D., & Liu, Y. (2018). Recurrent neural networks for multivariate time series with missing values. *Scientific Reports*, 8(1), 6085. <https://doi.org/10.1038/s41598-018-24271-9>
 - [114] Toye, A. A., Ahmad, A., Ibrahim, A., & Adamu, A. (2025). Benchmarking Missing Data Imputation Methods for Time-Series Healthcare Data. *Frontiers in Big Data*, 8, 1536374. <https://doi.org/10.3389/fdata.2025.1536374>
 - [115] Afkanpour, M., Daghistani, T., Mansourvar, M., ... (2024). Identify the most appropriate imputation method for extremely missing numerical data in Electronic Health Records for machine learning. *BMC Medical Informatics and Decision Making*, 24, 184. <https://doi.org/10.1186/s12911-024-02805-9>
 - [116] Junaid, K. P., et al. (2025). How much missing data is too much to impute for longitudinal health indicators. *Population Health Metrics*, 23(1), 5. <https://doi.org/10.1186/s12963-025-00364-2>
 - [117] Radosavljević, L., Đorđević, M., & Trajković, D. (2025). A generative model for evaluating missing data methods in large-scale epidemiological datasets. *BMC Medical Research Methodology*, 25, 25. <https://doi.org/10.1186/s12874-025-02487-4>
 - [118] Yoon, J., Jordon, J., & van der Schaar, M. (2020). GAIN: Missing data imputation using generative adversarial nets. *Proceedings of the International Conference on Machine Learning (ICML)*. <https://doi.org/10.48550/arXiv.1806.02920>
 - [119] Ais'sy, Widya Rohadatul and Firdaus, Firdaus and Tutuko, Bambang (2023). *Multivariate Imputation Tanda Vital Pasien Unit Perawatan Intensif Menggunakan Convolutional Neural Network*. <https://repository.unsri.ac.id/104475/>
 - [120] Dataset MIMIC-IV. Published: Oct. 11, 2024. Version: 3.1. <https://physionet.org/content/mimiciv/3.1/>