

BAB III

PELAKSANAAN PRO-STEP : ROAD TO CHAMPION

3.1 Kedudukan dan Koordinasi

Bagian ini menjelaskan posisi struktural penulis dalam pengerjaan proyek ilmiah "PEMODELAN PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE" dan mekanisme koordinasi yang diterapkan selama proyek berlangsung.

1) Kedudukan Antara Dosen Pembimbing Kompetisi dengan Kelompok / Individu Peserta Kompetisi Akademik / Beberapa pihak yang Berkepentingan

Proyek ini dilaksanakan dalam kerangka Program *Independent Project* (PRO-STEP: Road to Champion) Universitas Multimedia Nusantara (UMN) dengan jalur kompetisi Lomba Karya Tulis Ilmiah Nasional (LKTIN) CARBON 7.0. Dalam struktur tim, penulis (Marcellinus Gillbert Sunioto) memiliki kedudukan sebagai Ketua Tim.

Sebagai Ketua Tim, penulis bertanggung jawab penuh atas keseluruhan perencanaan, eksekusi teknis, validasi hasil, dan penyelesaian laporan akhir, serta bertindak sebagai penghubung utama antara tim, Dosen Pembimbing Internal, dan pihak penyelenggara lomba.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Pencatatan Rangkuman Mingguan Proses *PRO-STEP: Road To Champion Program*

Tabel 3. 1 Detail Pekerjaan yang Dilakukan PRO-STEP : Road to Champion Program

No.	Minggu	Proyek	Keterangan
1	1	Mencari Informasi lomba bersama dosen terkait dan penentuan topik	Mendapatkan Info lomba bersama Dosen dan team lain serta mendapatkan topik yang akan dibahas
2	2	Pembuatan Abstrak Laporan dan draft laporan	Membuat Abstrak bersama team dan melakukan submission untuk pengumpulan Abstrak Lomba
3	2	Pembuatan Laporan Full Paper	Membuat Laporan Full Paper dan melakukan submission setelah mendapat informasi lolos pemilihan Abstrak
4	3	Berkonsultasi bersama dosen dan pembuatan slide Presentasi serta design poster	Berkonsultasi bersama dosen mengenai laporan sekaligus membahas tentang PPT dan design Poster
5	4	Berangkat ke Jambi dan Technical Meeting	Kelompok berangkat ke Jambi dan melakukan Technical Meeting bersama Semua Team yang lolos seleksi
6	4	Presentasi Laporan	Melakukan Presentasi di Universitas Jambi bersama Team
7	4	Melakukan Field Trip	Melakukan Field Trip bersama Team TOP 10 lainnya ke Candi

No.	Minggu	Proyek	Keterangan
8	4	Pengumuman Hasil Lomba	Mendapatkan hasil lomba

Uraian Pelaksanaan Kerja Dalam *PRO-STEP : Road To Champion Program*

3.1.1 Proses Pelaksanaan

Proses pelaksanaan proyek ilmiah "PEMODELAN PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE" ini merupakan bagian integral dari Program PRO-STEP: Road to Champion Universitas Multimedia Nusantara (UMN) dengan target Lomba Karya Tulis Ilmiah Nasional (LKTIN) CARBON 7.0. Secara umum, pelaksanaan proyek mengikuti kerangka kerja metodologi *Cross-Industry Standard Process for Data Mining* (CRISP-DM) yang dimodifikasi untuk konteks kompetisi ilmiah, berfokus pada tahapan dari pemahaman masalah hingga implementasi dan evaluasi model.

3.1.1.1 Tahap 1

Penulis memulai tahap ini dengan membangun koordinasi yang intens bersama Dosen Koordinator serta tim mahasiswa PRO-STEP lainnya untuk memastikan bahwa arah pengerjaan proyek selaras dengan tujuan program. Tahap awal difokuskan pada Identifikasi Kompetisi, yaitu menentukan ajang ilmiah yang tepat untuk dijadikan target pelaksanaan proyek. Setelah melalui beberapa pertimbangan, tim memilih Lomba Karya Tulis Ilmiah Nasional (LKTIN) CARBON 7.0 yang diselenggarakan oleh Himpunan Mahasiswa Kimia Universitas Jambi (UNJA). Kompetisi ini dipandang relevan karena menekankan inovasi teknologi dan penelitian ilmiah, sehingga cocok dengan bidang yang ingin dikembangkan oleh penulis.

Setelah kompetisi ditetapkan, proses berlanjut pada Penentuan Topik, di mana penulis dan tim mendiskusikan isu apa yang memiliki urgensi di masyarakat sekaligus memungkinkan untuk dianalisis melalui pendekatan

teknologi. Dari berbagai alternatif, diputuskan bahwa penelitian akan berfokus pada penerapan Machine Learning di bidang kesehatan, khususnya prediksi penyakit jantung, karena penyakit jantung masih menjadi penyebab kematian tertinggi secara global dan termasuk isu krusial dalam konteks kesehatan nasional. Pemilihan topik ini juga mempertimbangkan ketersediaan data, peluang kontribusi ilmiah, serta kesesuaiannya dengan kemampuan teknis tim.

Tahap berikutnya adalah Pembentukan Tim Inti, di mana peran masing-masing anggota ditetapkan agar proses pengerjaan proyek berlangsung lebih terstruktur. Pada tahap ini, penulis mengambil peran utama sebagai penanggung jawab teknis pengembangan model dan penyusunan dokumen karya tulis ilmiah (KTI). Pembagian tugas ini bertujuan untuk memastikan setiap bagian proyek dapat dikerjakan secara fokus dan efisien.

Setelah topik dikonfirmasi dan disetujui oleh Dosen Pembimbing, penulis memasuki tahap Studi Literatur Awal dan Pengumpulan Referensi. Proses ini melibatkan pencarian terarah terhadap jurnal ilmiah nasional dan internasional yang membahas prediksi penyakit jantung menggunakan algoritma Machine Learning. Penulis menelaah metode-metode yang umum digunakan, seperti Logistic Regression, Random Forest, KNN, dan Support Vector Machine (SVM). Dari hasil studi literatur tersebut, SVM dipilih sebagai algoritma utama karena dikenal memiliki performa baik pada data klasifikasi berukuran kecil hingga menengah, serta mampu bekerja optimal dalam pemisahan kelas yang kompleks. Tahap ini sekaligus menjadi fondasi dalam penentuan metodologi awal yang akan digunakan dalam penyusunan laporan dan pengembangan model pada proses kompetisi.

3.1.1.2 Tahap 2

Tahap ini menjadi salah satu *milestone* awal yang paling penting dalam keseluruhan proses kompetisi. Setelah topik penelitian dan arah metodologi ditetapkan, penulis bersama tim mulai memasuki tahap penyusunan Abstrak, yang merupakan dokumen pertama yang dinilai oleh panitia untuk menentukan kelayakan proposal kami. Penyusunan abstrak dilakukan melalui koordinasi aktif antar anggota tim, termasuk diskusi mengenai poin-poin utama yang harus ditonjolkan, seperti urgensi masalah penyakit jantung, relevansi penggunaan algoritma Machine Learning, serta kontribusi yang ingin diberikan melalui penelitian ini.

Format abstrak disusun mengikuti ketentuan resmi LKTIN CARBON 7.0, yaitu maksimal 250 kata dan harus mencakup elemen penting seperti latar belakang penelitian, tujuan studi, gambaran metode yang akan digunakan (termasuk pemilihan SVM sebagai algoritma inti), data yang direncanakan untuk diolah, serta perkiraan hasil dan manfaat penelitian. Proses ini membutuhkan beberapa kali revisi untuk memastikan abstrak tampil ringkas, jelas, dan mampu merepresentasikan keseluruhan rencana penelitian dengan baik.

Tahap Pembuatan Abstrak dan Submission ini juga menjadi momen penting karena menandai pengajuan resmi proposal proyek kepada panitia kompetisi. Dengan terkirimnya abstrak, tim memasuki fase verifikasi dan seleksi awal, sekaligus menegaskan bahwa proyek yang sedang dikerjakan sudah berada pada jalur kompetisi yang jelas dan terarah.

3.1.1.3 Tahap 3

Setelah tim menerima informasi lolos seleksi Abstrak, kami langsung melanjutkan dengan menyusun Laporan *Full Paper*. Penulisan draf laporan ini kami kerjakan bersamaan dengan tahapan implementasi teknis. Hal ini penting untuk memastikan hasil riset teknis dapat langsung

terintegrasi dan dibahas dalam dokumen laporan. Untuk memberikan bukti dan gambaran visual mengenai tahap awal penyusunan laporan serta proses pengiriman berkas, kami menyajikan beberapa tangkapan layar di bawah ini.

1. Proses Pembuatan Latar Belakang

Pada tahap penyusunan latar belakang, tim memulai proses dengan mengidentifikasi permasalahan utama yang ingin diangkat, yaitu tingginya angka kematian akibat penyakit jantung yang masih menjadi salah satu penyebab kematian terbesar di dunia maupun di Indonesia. Untuk memperkuat argumentasi, tim mengumpulkan berbagai data dan fakta pendukung dari sumber-sumber terpercaya seperti jurnal ilmiah internasional, laporan statistik kesehatan dari WHO, artikel medis yang terakreditasi, hingga publikasi penelitian terbaru yang relevan. Informasi tersebut kemudian dianalisis untuk memahami pola, faktor risiko, serta urgensi kebutuhan akan inovasi di bidang deteksi dini penyakit jantung.

Proses pengolahan data ini tidak hanya berfokus pada pemaparan masalah, tetapi juga diarahkan untuk membangun narasi yang koheren dan logis mengenai pentingnya penggunaan teknologi modern, khususnya machine learning, dalam membantu proses diagnosis. Tim secara sistematis merumuskan bagaimana metode prediktif berbasis algoritma dapat menawarkan kecepatan, efisiensi, dan tingkat akurasi yang sulit dicapai melalui metode manual konvensional. Dengan pendekatan tersebut, latar belakang akhirnya disusun dengan tujuan menunjukkan relevansi penelitian serta memberikan gambaran bahwa proyek pengembangan model prediksi penyakit jantung ini memiliki manfaat nyata, baik dari sisi akademik maupun sosial, terutama dalam upaya deteksi dini dan pencegahan risiko penyakit jantung.

2. Proses pembuatan Studi Literatur

Pada tahap penyusunan studi literatur, tim melakukan pencarian dan pengumpulan referensi secara sistematis untuk memahami landasan teori yang diperlukan dalam penelitian. Proses ini dimulai dengan menelusuri berbagai jurnal ilmiah, artikel penelitian, serta publikasi akademik yang membahas algoritma Machine Learning, metode preprocessing data kesehatan, dan penelitian terdahulu terkait prediksi penyakit jantung. Sumber-sumber referensi utama diambil dari database kredibel seperti IEEE Xplore, ScienceDirect, dan Frontiers, sehingga informasi yang digunakan memiliki tingkat validitas dan keandalan yang tinggi.

Setiap jurnal yang ditemukan kemudian dianalisis untuk melihat metode yang digunakan, cara pengolahan data, tingkat akurasi yang dicapai, serta kekuatan dan kelemahan pendekatannya. Melalui proses perbandingan ini, tim mengidentifikasi gap penelitian yang relevan, yaitu kebutuhan akan model prediksi yang mampu mencapai akurasi lebih tinggi dengan memanfaatkan algoritma Support Vector Machine (SVM) pada dataset yang sesuai. Temuan-temuan dari literatur tersebut menjadi dasar penting dalam merumuskan arah metodologi dan justifikasi pemilihan algoritma.

Seluruh teori, konsep pendukung, dan hasil penelitian sebelumnya yang memiliki relevansi langsung dengan topik proyek kemudian dirangkum secara terstruktur dalam bagian tinjauan pustaka. Penyusunan ini bertujuan untuk memastikan bahwa pengembangan model prediksi penyakit jantung yang dilakukan memiliki basis ilmiah yang kuat dan terhubung secara logis dengan penelitian-penelitian yang telah ada sebelumnya.

3. Proses Pembuatan Metode Penelitian

Pada tahap penyusunan metode penelitian, tim mengembangkan alur kerja yang sistematis sesuai proses pemodelan Machine Learning yang digunakan dalam proyek ini. Metodologi ini disusun untuk memastikan bahwa setiap langkah yang dilakukan dapat dipertanggungjawabkan secara ilmiah dan selaras dengan standar penelitian data science. Tahapan pertama dimulai dari pemilihan dataset, yaitu Heart Failure Prediction Dataset dari Kaggle, yang dipilih karena memiliki struktur data yang lengkap dan relevan dengan kebutuhan analisis prediksi penyakit jantung.

Setelah dataset ditentukan, tim melanjutkan ke proses pembersihan data (data cleaning), termasuk pemeriksaan nilai kosong (missing values), eliminasi inkonsistensi, serta peninjauan kelayakan tiap fitur. Tahap selanjutnya adalah pengecekan outlier menggunakan metode statistik seperti Interquartile Range (IQR), untuk memastikan tidak adanya nilai ekstrim yang berpotensi mengganggu proses pelatihan model. Apabila ditemukan outlier, data tersebut dianalisis apakah perlu dihapus atau tetap dipertahankan tergantung konteks medisnya.

Untuk meningkatkan performa model, dilakukan standarisasi fitur numerik menggunakan StandardScaler. Langkah ini penting karena algoritma Support Vector Machine (SVM) sangat bergantung pada skala data, sehingga penyetaraan skala dapat mencegah bias terhadap fitur tertentu. Setelah data distandarisasi, dataset dibagi ke dalam dua subset utama, yaitu data latih (training set) dan data uji (testing set), guna mengevaluasi kemampuan generalisasi model terhadap data baru.

Proses inti dari metode penelitian selanjutnya adalah pelatihan model menggunakan algoritma SVM dengan kernel Radial Basis Function (RBF), yang dipilih karena kemampuannya dalam menangani data non-linear dan menghasilkan batas keputusan yang optimal. Setelah model dilatih, dilakukan evaluasi performa dengan beberapa metrik penting, yaitu akurasi, precision, recall, dan f1-score, yang memberikan gambaran komprehensif mengenai kemampuan model dalam memprediksi penyakit jantung secara tepat.

Untuk memperjelas keseluruhan proses metodologi, seluruh tahapan tersebut divisualisasikan melalui diagram alur (flowchart). Diagram ini berfungsi sebagai representasi visual yang membantu pembaca memahami alur teknis penelitian secara lebih terstruktur dan mudah dipahami.

4. Proses Pembuatan Hasil dan Pembahasan

a. Pengumpulan Data

Sumber data sekunder yang digunakan dalam penelitian ini adalah "Heart Failure Prediction Dataset" yang diperoleh dari platform Kaggle. Data tersebut disajikan dalam format berkas *.csv*. Seluruh tahapan kerja, mulai dari akuisisi data, proses persiapan data (*preprocessing*), implementasi model *Support Vector Machine* (SVM), hingga analisis dan evaluasi kinerja model, dilaksanakan secara komprehensif dalam lingkungan komputasi Jupyter Notebook.

```
heart.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 918 entries, 0 to 917  
Data columns (total 12 columns):  
#      Column              Non-Null Count  Dtype  
---  -  
0     Age                   918 non-null   int64  
1     Sex                    918 non-null   object  
2     ChestPainType          918 non-null   object  
3     RestingBP              918 non-null   int64  
4     Cholesterol             918 non-null   int64  
5     FastingBS              918 non-null   int64  
6     RestingECG             918 non-null   object  
7     MaxHR                  918 non-null   int64  
8     ExerciseAngina         918 non-null   object  
9     Oldpeak                918 non-null   float64  
10    ST_Slope               918 non-null   object  
11    HeartDisease           918 non-null   int64  
dtypes: float64(1), int64(6), object(5)
```

Gambar 3. 1 Fitur – fitur yang terdapat di dalam Dataset

Gambar 3.1 menunjukkan dataset yang digunakan dalam penelitian ini, yaitu *Heart Failure Prediction Dataset* yang diperoleh dari platform Kaggle. Dataset ini terdiri dari 918 catatan pasien dengan total 12 fitur yang merepresentasikan informasi demografis dan klinis terkait risiko penyakit jantung. Fitur-fitur tersebut mencakup data dasar seperti usia dan jenis kelamin, serta atribut medis yang umum digunakan dalam pemeriksaan kardiovaskular, seperti *ChestPainType*, *RestingBP*, *Cholesterol*, *FastingBS*, *RestingECG*, *MaxHR*, *ExerciseAngina*, *Oldpeak*, dan *ST_Slope*. Kombinasi fitur-fitur tersebut memberikan gambaran kondisi kesehatan pasien secara menyeluruh dan relevan untuk proses analisis serta pemodelan prediksi penyakit jantung. Seluruh atribut digunakan untuk memprediksi variabel target *HeartDisease*, yaitu label biner yang menunjukkan ada atau tidaknya diagnosis penyakit jantung pada pasien. Dataset ini bersifat *cross-sectional* dan tidak mencantumkan informasi rentang tahun pengambilan data, karena data yang digunakan merupakan kompilasi data klinis pasien pada satu titik

pemeriksaan dan tidak ditujukan untuk analisis berbasis waktu. Dengan cakupan fitur yang lengkap dan relevan, dataset ini menjadi dasar yang kuat untuk membangun model Machine Learning berbasis klasifikasi yang diharapkan mampu menghasilkan performa prediksi yang akurat dan konsisten.

b. Pre Processing Data

Kelompok kami melaksanakan tahap pra-pemrosesan data sebagai langkah awal untuk mempersiapkan dataset sebelum masuk ke proses pelatihan model Machine Learning. Tahap ini dilakukan untuk memastikan bahwa data berada dalam kondisi yang bersih, konsisten, dan siap digunakan, karena kualitas data sangat berpengaruh terhadap performa model klasifikasi. Pada proses ini, kelompok meninjau kemungkinan adanya *noise*, nilai yang tidak relevan, data yang tidak lengkap, hingga inkonsistensi antar fitur yang dapat memengaruhi akurasi prediksi. Selain itu, dilakukan pula pengecekan terhadap tipe data, penanganan nilai ekstrem, dan penyesuaian format agar seluruh variabel yang digunakan sesuai dengan kebutuhan algoritma. Pra-pemrosesan menjadi langkah penting karena model klasifikasi seperti SVM membutuhkan input yang terstruktur dan terstandarisasi agar dapat bekerja secara optimal.

```
#mencari missing value
heart.isna().sum()
```

```
Age          0
Sex          0
ChestPainType 0
RestingBP    0
Cholesterol  0
FastingBS    0
RestingECG   0
MaxHR        0
ExerciseAngina 0
Oldpeak      0
ST_Slope     0
HeartDisease 0
dtype: int64
```

Gambar 3. 2 Pencarian missing value

Pada tahap pembersihan data, kelompok terlebih dahulu melakukan pemeriksaan terhadap nilai yang hilang (*missing values*) dengan menggunakan fungsi `isna().sum()` pada dataframe, seperti yang ditampilkan pada Gambar 3.2. Pemeriksaan ini dilakukan untuk memastikan bahwa tidak ada kolom yang memiliki data kosong yang dapat mengganggu proses pelatihan model. Berdasarkan hasil pengecekan, seluruh kolom pada dataset terkonfirmasi tidak memiliki *missing values*, sehingga dataset dapat dianggap bersih dari kekosongan data dan tidak memerlukan proses imputasi tambahan. Setelah memastikan bahwa dataset sepenuhnya lengkap, proses pembersihan data kemudian dilanjutkan ke tahap berikutnya, yaitu identifikasi dan analisis *outlier*, untuk memastikan bahwa tidak terdapat nilai ekstrem yang berpotensi memengaruhi distribusi data atau menurunkan performa model klasifikasi yang akan dibangun.

```
#melihat outlier

Q1 = heart['Age'].quantile(0.25)
Q3 = heart['Age'].quantile(0.75)
IQR = Q3 - Q1

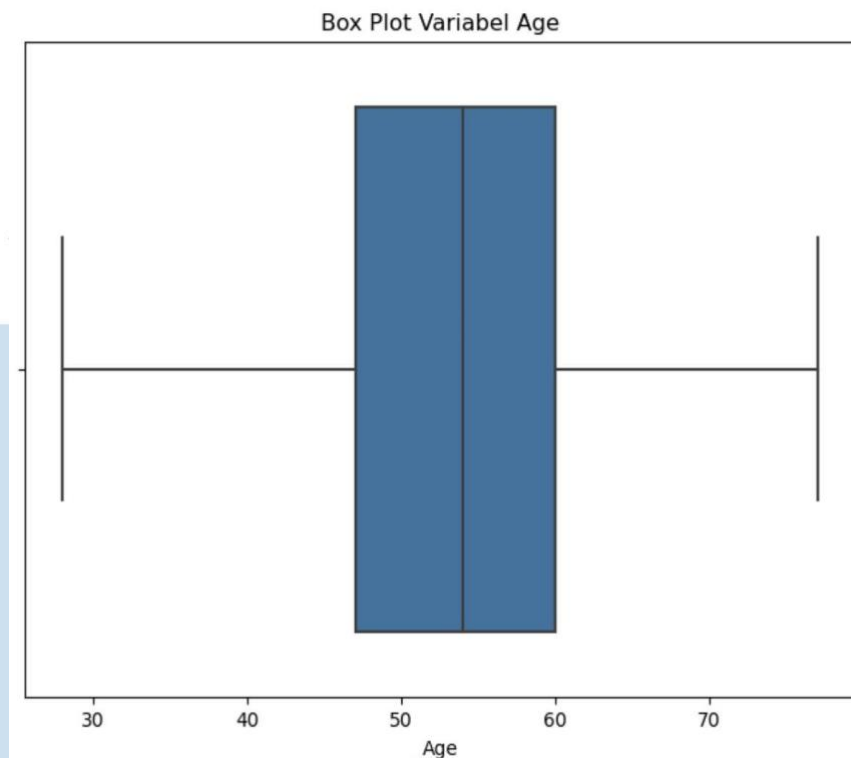
lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

outliers = heart[(heart['Age'] < lower_bound) | (heart['Age'] > upper_bound)]
print("Jumlah outlier:", len(outliers))
print(outliers)

Jumlah outlier: 0
Empty DataFrame
Columns: [Age, Sex, ChestPainType, RestingBP, Cholesterol, FastingBS, RestingECG, MaxHR, ExerciseAngina, Oldpeak, S
T_Slope, HeartDisease]
Index: []
```

Gambar 3. 3 Melihat Outlier

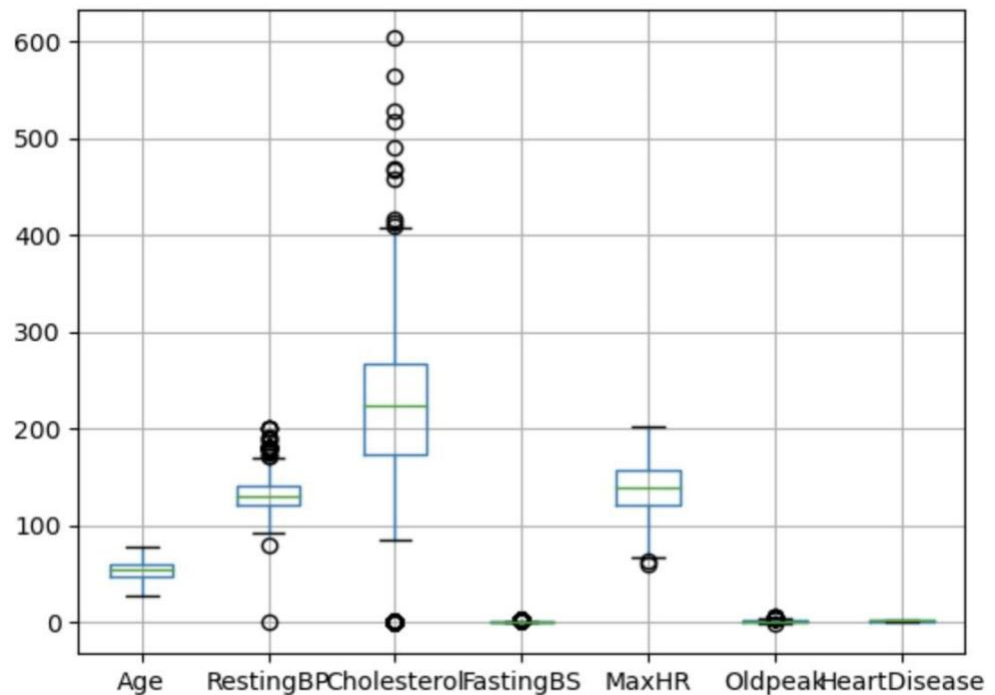
Pemeriksaan terhadap keberadaan *outlier* kemudian dilakukan untuk memastikan bahwa distribusi data pada setiap fitur berada dalam rentang yang wajar dan tidak terdapat nilai ekstrem yang dapat memengaruhi performa model. Berdasarkan hasil visualisasi pada Gambar 3.3, dapat dilihat bahwa kolom Age tidak memiliki *outlier* sama sekali, dengan jumlah nilai ekstrem yang terdeteksi adalah 0. Hal ini menunjukkan bahwa seluruh data usia dalam dataset berada dalam batas bawah dan batas atas yang dapat diterima secara statistik, sehingga fitur tersebut dapat dianggap stabil dan tidak memerlukan proses penanganan tambahan. Kondisi ini juga menjadi indikator bahwa dataset memiliki kualitas distribusi yang baik, sehingga proses analisis dapat dilanjutkan tanpa perlu menghapus atau menyesuaikan nilai tertentu. Dengan tidak adanya *outlier* pada fitur ini, dataset kemudian dinilai sudah memenuhi syarat untuk melangkah ke tahap berikutnya, yaitu proses pembagian data menjadi training set dan testing set untuk keperluan implementasi model klasifikasi.



Gambar 3. 4 Boxplot Variable Age

Gambar 3.4 menampilkan analisis distribusi data usia menggunakan boxplot yang memberikan gambaran visual mengenai persebaran usia pada seluruh sampel pasien dalam dataset. Dari visualisasi tersebut terlihat bahwa median usia berada di sekitar 55 tahun, yang menandakan bahwa setengah dari total populasi dataset berusia di bawah angka tersebut dan setengahnya lagi berada di atasnya. Rentang interkuartil (Interquartile Range/IQR), yang direpresentasikan oleh bagian kotak pada boxplot dan mencakup 50% data tengah, berada antara kuartil pertama (Q1) sekitar 48 tahun hingga kuartil ketiga (Q3) sekitar 60 tahun, sehingga menunjukkan bahwa sebagian besar pasien berada dalam rentang usia paruh baya hingga awal lansia. Sementara itu, garis whiskers pada boxplot menggambarkan sebaran data secara keseluruhan, yaitu dari usia minimum sekitar 30 tahun hingga usia maksimum sekitar 75 tahun, tanpa adanya nilai ekstrem yang menyimpang. Secara

keseluruhan, distribusi ini menunjukkan bahwa dataset memiliki variasi usia yang cukup seimbang dan mencerminkan populasi yang relevan untuk studi mengenai risiko penyakit jantung, sehingga layak digunakan dalam proses pelatihan model klasifikasi.



Gambar 3. 5 Boxplot mengenai Distribusi dan Karakteristik setiap Variabel

Visualisasi distribusi data menggunakan boxplot pada Gambar 3.5 memberikan gambaran menyeluruh mengenai pola persebaran serta karakteristik dari setiap variabel yang digunakan dalam penelitian. Secara umum, boxplot menunjukkan bahwa terdapat variasi yang cukup signifikan pada beberapa fitur utama. Pada variabel Age, rentang usia partisipan terlihat berada di antara 30 hingga 70 tahun, dengan sejumlah *outlier* pada kedua ujung rentang tersebut yang menunjukkan adanya pasien berusia lebih muda maupun lebih tua dari mayoritas sampel. Fitur RestingBP memperlihatkan adanya *outlier* di atas nilai maksimal, yang mengindikasikan keberadaan pasien dengan tekanan darah istirahat yang sangat tinggi dibandingkan kelompok lainnya. Sementara itu, fitur Cholesterol memiliki jumlah *outlier* yang cukup

banyak, terutama pada bagian atas distribusi, yang menandakan adanya pasien dengan kadar kolesterol ekstrem, bahkan hingga melebihi angka 400. Variabel MaxHR atau detak jantung maksimum tampak berada dalam rentang yang relatif moderat, namun tetap menunjukkan beberapa nilai *outlier* di bagian bawah distribusi. Selain itu, fitur seperti FastingBS, Oldpeak, dan HeartDisease cenderung memiliki distribusi nilai yang lebih rendah tetapi tetap memunculkan sejumlah *outlier* yang terdeteksi. Secara keseluruhan, hasil visualisasi ini menunjukkan bahwa dataset mengandung karakteristik variasi nilai yang cukup beragam, sehingga tahap pra-pemrosesan seperti identifikasi dan penanganan *outlier* menjadi penting untuk memastikan kualitas data yang optimal sebelum memasuki proses pelatihan model.

```
#standardization
from sklearn.preprocessing import StandardScaler

scaler = StandardScaler()
numerical_features = ['Age', 'RestingBP', 'Cholesterol', 'MaxHR', 'Oldpeak', 'FastingBS', 'HeartDisease']
heart[numerical_features] = scaler.fit_transform(heart[numerical_features])
print(heart.head())
```

	Age	Sex	ChestPainType	RestingBP	Cholesterol	FastingBS	Restin
gECG \							
0	-1.433140	M	ATA	0.410909	0.825070	0	No
1	-0.478484	F	NAP	1.491752	-0.171961	0	No
2	-1.751359	M	ATA	-0.129513	0.770188	0	
3	-0.584556	F	ASY	0.302825	0.139040	0	No
4	0.051881	M	NAP	0.951331	-0.034755	0	No
	MaxHR	ExerciseAngina	Oldpeak	ST_Slope	HeartDisease		
0	1.382928	N	-0.832432	Up	0		
1	0.754157	N	0.105664	Flat	1		
2	-1.525138	N	-0.832432	Up	0		
3	-1.132156	Y	0.574711	Flat	1		
4	-0.581981	N	-0.832432	Up	0		

Gambar 3. 6 Standarisasi Fitur

Gambar 3.6 menggambarkan Tahap selanjutnya dalam proses pra-pemrosesan adalah melakukan standarisasi pada fitur numerik menggunakan fungsi `StandardScaler` dari pustaka `scikit-learn`. Standarisasi ini diterapkan pada beberapa variabel numerik utama, yaitu Age, RestingBP, Cholesterol, MaxHR, dan Oldpeak, dengan tujuan mentransformasikan nilai setiap fitur agar memiliki rata-rata (mean) sebesar 0 dan standar deviasi (standard deviation) sebesar 1. Transformasi ini sangat penting karena algoritma Support Vector Machine (SVM) cukup sensitif terhadap perbedaan skala antar fitur; jika tidak dilakukan standarisasi, fitur dengan nilai yang lebih besar dapat mendominasi proses pelatihan dan memengaruhi pembentukan hyperplane secara tidak proporsional. Dengan dilakukan standarisasi, seluruh fitur numerik berada pada skala yang setara sehingga model dapat belajar secara lebih seimbang. Sementara itu, fitur kategorikal seperti Sex, ChestPainType, ExerciseAngina, ST_Slope, dan HeartDisease dibiarkan apa adanya karena tidak memerlukan proses standarisasi dan tetap digunakan sesuai bentuk kategorinya untuk mendukung proses klasifikasi.

c. Klasifikasi Data

Setelah proses standarisasi fitur numerik selesai dilakukan, langkah berikutnya adalah membagi dataset menjadi dua bagian utama, yaitu training set dan testing set. Pembagian ini memiliki peran penting dalam proses pembangunan model Machine Learning karena memungkinkan evaluasi performa model secara objektif. Data pelatihan digunakan untuk melatih algoritma agar dapat mengenali pola dan hubungan antar fitur, sedangkan data pengujian digunakan untuk menilai seberapa baik model mampu melakukan prediksi pada data yang belum pernah ditemuinya sebelumnya. Dengan cara ini, penilaian tidak hanya bergantung pada kemampuan model mengingat pola dari

data latih, tetapi juga pada kemampuan generalisasinya di situasi nyata. Pemisahan ini juga berfungsi untuk mengurangi risiko overfitting, yaitu kondisi ketika model tampil sangat baik pada data latih namun gagal memprediksi data baru dengan akurat. Oleh karena itu, pembagian dataset menjadi training dan testing set menjadi langkah penting untuk memastikan bahwa model yang dibangun tidak hanya kuat secara teoritis, tetapi juga dapat diterapkan secara efektif di dunia nyata.

```
from sklearn.svm import SVC
from sklearn.preprocessing import StandardScaler
from sklearn.pipeline import make_pipeline

scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.2, random_state=42)

svm_classifier = make_pipeline(StandardScaler(), SVC(kernel='rbf', random_state=42))
svm_classifier.fit(X_train, y_train)

y_pred_svm = svm_classifier.predict(X_test)

accuracy_svm = accuracy_score(y_test, y_pred_svm)
print("Accuracy:", accuracy_svm)
print("\nClassification Report:")
print(classification_report(y_test, y_pred_svm))
```

Accuracy: 0.8913043478260869

Classification Report:

	precision	recall	f1-score	support
0	0.86	0.88	0.87	77
1	0.91	0.90	0.91	107
accuracy			0.89	184
macro avg	0.89	0.89	0.89	184
weighted avg	0.89	0.89	0.89	184

Gambar 3. 7 Proses Klasifikasi menggunakan SVM

Gambar 3.7 menunjukkan proses klasifikasi menggunakan Support Vector Machine (SVM) yang diawali dengan pembentukan *pipeline* yang

mengintegrasikan *StandardScaler* dan model *Support Vector Classifier* (SVC) dengan fungsi kernel. Dataset yang telah melalui proses standarisasi kemudian dibagi menjadi data training dan data testing dengan rasio 80:20. Pembagian ini dipilih untuk memberikan porsi data yang cukup besar pada proses pelatihan model, sehingga model dapat mempelajari pola data secara optimal, sekaligus tetap menyediakan data pengujian yang memadai untuk mengevaluasi performa model secara objektif. Rasio 80:20 juga merupakan praktik yang umum digunakan dalam penelitian Machine Learning, khususnya pada dataset dengan jumlah data yang terbatas. Proses pembagian data dilakukan dengan menggunakan parameter *random_state* = 42 untuk memastikan hasil eksperimen dapat direproduksi.

Setelah pelatihan menggunakan data training, model diuji menggunakan data testing (*X_test*), dan tingkat akurasi yang dicapai adalah 0.89 (atau 89%). Laporan klasifikasi menunjukkan kinerja yang sangat baik dan seimbang pada kedua kelas (0 dan 1):

- A. Untuk Kelas '0', nilai *precision*, *recall*, dan *f1-score* masing-masing adalah 0.86, 0.88, dan 0.87.
- B. Untuk Kelas '1', nilai *precision*, *recall*, dan *f1-score* masing-masing adalah 0.91, 0.90, dan 0.91.

Secara keseluruhan, model SVM berhasil mencapai akurasi total 0.89, dengan nilai rata-rata makro dan tertimbang untuk semua metrik utama juga berada pada angka 0.89. Hasil ini mengonfirmasi bahwa model SVM memiliki performa yang sangat baik dan stabil dalam mengklasifikasikan data risiko penyakit jantung.

Setelah model berhasil training dan testing, seluruh output berupa tabel akurasi, classification report, confusion matrix, dan hasil analisis distribusi data divisualisasikan dalam bentuk grafik dan gambar. Hasil performa SVM yang mencapai akurasi 89% dianalisis dan dibandingkan dengan penelitian sebelumnya untuk menunjukkan efektivitas model. Pembahasan dilakukan

untuk menjelaskan faktor-faktor risiko dominan yang memengaruhi prediksi. Selain itu, bagian ini juga menjadi dasar bagi penarikan kesimpulan dan rekomendasi pada bab selanjutnya.

5. Pembuatan Kesimpulan dan Saran

Pada tahap akhir penyusunan laporan, tim menyusun bagian kesimpulan berdasarkan seluruh hasil analisis yang diperoleh dari eksperimen model prediksi penyakit jantung menggunakan algoritma Support Vector Machine (SVM). Kesimpulan dirumuskan dengan mengacu pada tujuan utama penelitian, yaitu untuk mengetahui sejauh mana model dapat memprediksi risiko penyakit jantung secara akurat. Oleh karena itu, tim merangkum temuan inti berupa tingkat akurasi model sebesar 89%, fitur medis yang memiliki kontribusi paling besar dalam prediksi, serta relevansi solusi yang ditawarkan terhadap peningkatan deteksi dini penyakit jantung di dunia medis.

3.1.1.4 Tahap 4

Tahap keempat dari pelaksanaan proyek PRO-STEP ini dipusatkan pada kegiatan konsultasi dan verifikasi komprehensif bersama dosen pembimbing, yang merupakan mekanisme esensial untuk memvalidasi integritas ilmiah dan metodologis dari seluruh penelitian, sekaligus mempersiapkan persiapan materi presentasi dan poster untuk ajang LKTIN CARBON 7.0. Sesi konsultasi ini melibatkan peninjauan mendalam terhadap draft laporan akhir, memastikan konsistensi antara hasil eksperimen, khususnya kinerja optimal model *Support Vector Machine* (SVM) yang telah melalui *hyperparameter tuning ekstensif*, dengan narasi akademis. Selanjutnya, proses difokuskan pada perancangan sepuluh (10) *slide Power Point Presentation* (PPT) yang terstruktur, dimulai dari penyajian urgensi Latar Belakang mengenai krisis Penyakit Kardiovaskular dan kebutuhan solusi non-invasif, diikuti dengan perumusan Tujuan Penelitian yang spesifik dan penetapan Dataset dan

Tools yang digunakan (Python, Scikit-learn). Alur narasi presentasi kemudian bergerak melalui Alur Penelitian yang terinci, memberikan landasan teori mengenai Algoritma SVM, sebelum mencapai puncak pada Hasil Penelitian yang menampilkan perbandingan performa *baseline* dan hasil optimalisasi. Materi presentasi ditutup dengan Visualisasi matrik evaluasi seperti *Confusion Matrix* dan kurva ROC, serta Kesimpulan yang merangkum pencapaian dan Implikasi ke Depan dari model prediksi yang sukses dikembangkan. Secara paralel, dirancang pula Desain Poster Ilmiah sebagai ringkasan visual satu halaman yang menarik, berfokus pada kejelasan diagram alir metodologi dan keunggulan visualisasi hasil, memastikan materi yang disajikan tidak hanya informatif tetapi juga persuasif dan siap untuk menghadapi penilaian juri.

3.1.1.5 Tahap 5

Tahap kelima dari rangkaian kegiatan PRO-STEP merupakan fase implementasi mobilitas dan persiapan akhir di lokasi kompetisi. Tepat pada tanggal 6 November 2025 (sehari sebelum presentasi final), tim peneliti memulai perjalanan dari Jakarta menuju Jambi. Perjalanan udara dipilih sebagai moda transportasi utama untuk efisiensi waktu, di mana tim bertolak dari Bandar Udara Internasional Soekarno-Hatta (CGK) pada pukul 10.00 WIB. Penerbangan yang relatif singkat, memakan waktu kurang lebih satu (1) jam, membawa tim mendarat dengan aman di Bandar Udara Sultan Thaha Syaifuddin (DJB), Jambi. Begitu tiba, tim segera disambut dengan hangat oleh panitia penyelenggara dari Universitas Jambi (UNJA), yang telah menjadwalkan penjemputan terorganisir untuk memfasilitasi kelancaran transisi tim menuju akomodasi yang telah disediakan. Setelah proses *check-in* dan alokasi waktu singkat untuk pemulihan energi dan adaptasi lingkungan, fokus tim dialihkan pada agenda malam hari. Tepat pada pukul 20.00 WIB, seluruh finalis diwajibkan menghadiri Technical Meeting (TM), sebuah sesi krusial yang berfungsi sebagai platform koordinasi terakhir sebelum hari H presentasi. Dalam sesi TM yang dilaksanakan secara

formal namun informatif ini, panitia membahas secara mendetail seluruh Sistematika Presentasi, mencakup alokasi waktu yang ketat untuk pemaparan dan sesi tanya jawab, urutan penampilan (*running order*) tim finalis, serta berbagai Tata Tertib yang berlaku selama sesi presentasi dan penjurian berlangsung, termasuk kode etik berpakaian dan prosedur penanganan kendala teknis, memastikan setiap tim memiliki pemahaman yang seragam dan memadai mengenai standar pelaksanaan kompetisi.

3.1.1.6 Tahap 6

Tahap keenam merupakan puncak dari seluruh rangkaian kegiatan PRO-STEP, yaitu pelaksanaan Presentasi Final Lomba Karya Tulis Ilmiah Nasional (LKTIN) CARBON 7.0 yang bertempat di Universitas Jambi (UNJA) pada tanggal 7 November 2025. Acara dimulai secara resmi tepat pukul 08.00 WIB, diawali dengan pembukaan seremonial yang khidmat dan tertata, yang mencerminkan standar tinggi penyelenggaraan kompetisi ilmiah. Sesi pembukaan dimulai dengan menyanyikan Lagu Kebangsaan Indonesia Raya sebagai penegasan semangat nasionalisme dan integritas keilmuan, dilanjutkan dengan lantunan Mars Universitas Jambi yang menyuarakan semangat dan visi institusi tuan rumah—sebuah momen penting untuk menghormati penyelenggara dan suasana akademik. Setelah sesi seremonial, yang juga mencakup kata pengantar resmi dari perwakilan pimpinan universitas dan pembacaan doa pembuka, dimulailah sesi presentasi yang sangat dinantikan. Setiap tim finalis ditantang untuk menyajikan hasil riset mereka di hadapan dewan juri yang terdiri dari tiga (3) akademisi terkemuka di bidang *data science* dan kesehatan. Berdasarkan urutan penampilan (*running order*) yang telah ditetapkan pada *Technical Meeting* malam sebelumnya, kelompok kami memperoleh kesempatan sebagai penyaji urutan ketiga, sebuah posisi yang memberikan waktu yang cukup untuk observasi terhadap dinamika penjurian tetapi tetap

mempertahankan momentum energi dan fokus.



Gambar 3. 8 Tampilan Presentasi Kelompok

Gambar 3.8 menunjukkan waktu saat maju ke podium, tim kami memanfaatkan seluruh alokasi waktu presentasi selama lima belas (15) menit secara optimal, menyajikan sepuluh *slide*PPT yang telah divalidasi dengan lancar dan meyakinkan, memaparkan secara sistematis metodologi optimalisasi *Support Vector Machine* (SVM) dan hasil akurasi prediksi penyakit jantung. Sesi krusial berikutnya adalah Tanya Jawab dengan durasi yang sama, yaitu lima belas (15) menit, di mana dewan juri melontarkan serangkaian pertanyaan mendalam yang menantang, mulai dari justifikasi pemilihan arsitektur model, validasi metrik kinerja, hingga implikasi praktis dan potensi keberlanjutan dari sistem prediksi yang dikembangkan. Berkat persiapan konsultasi yang matang dan pemahaman tim yang menyeluruh, kami mampu merespons setiap pertanyaan dari

ketiga juri—meliputi aspek keilmuan, teknis, dan implementasi—dengan terstruktur, lancar, dan lugas. Keberhasilan dalam sesi tanya jawab ini menegaskan penguasaan penuh terhadap materi penelitian yang telah dilaksanakan, sehingga menutup Tahap Presentasi Final dengan *performance* yang memuaskan dan memberikan *impact* ilmiah yang kuat di hadapan dewan penilai.

3.1.1.7 Tahap 7

Tahap ketujuh merupakan agenda non-akademis namun penting, yaitu pelaksanaan Field Trip bagi sepuluh (10) tim finalis LKTIN CARBON 7.0 yang diselenggarakan pada tanggal 8 November 2025, setelah seluruh rangkaian presentasi final selesai. Kegiatan ini bertujuan untuk memberikan kesempatan kepada para peserta dari berbagai daerah untuk berinteraksi, bertukar wawasan, dan mengeksplorasi kekayaan warisan budaya serta sejarah Provinsi Jambi, sehingga memperkaya pengalaman di luar konteks kompetisi. Destinasi utama yang telah dijadwalkan adalah kompleks bersejarah Candi Muaro Jambi, sebuah kawasan cagar budaya yang diyakini sebagai peninggalan Kerajaan Sriwijaya dan merupakan kompleks candi terluas di Asia Tenggara. Kunjungan ke situs bersejarah ini menyediakan wawasan berharga mengenai peradaban kuno yang pernah berkembang di wilayah tersebut. Meskipun agenda awal mencakup rencana untuk mengunjungi beberapa masjid terkenal di sekitar Kota Jambi sebagai bagian dari eksplorasi arsitektur dan religi lokal, sayangnya jadwal kunjungan tersebut tidak dapat direalisasikan karena waktu operasional destinasi yang belum dibuka untuk umum, sehingga membatasi rangkaian eksplorasi tim pada hari itu. Oleh karena kendala waktu tersebut, fokus *field trip* secara penuh dialihkan ke Candi Muaro Jambi, memungkinkan tim untuk menyerap informasi historis dan menikmati keindahan arsitektur kuno secara maksimal. Kegiatan *field trip* ini secara efektif berfungsi sebagai kegiatan rekreatif sekaligus edukatif yang mempererat hubungan antar finalis dan memberikan perspektif baru mengenai signifikansi konteks geografis dan

historis Jambi dalam kancah nasional.

3.1.1.8 Tahap 8

Tahap kedelapan menandai penutupan resmi seluruh rangkaian kegiatan LKTIN CARBON 7.0 yang dilaksanakan di Kampus Universitas Jambi (UNJA) pada hari berikutnya, dimulai tepat pukul 08.00 WIB. Acara penutupan ini diawali dengan sesi seremonial yang serupa dengan pembukaan, yaitu menyanyikan Lagu Kebangsaan Jakarta Raya dan dilanjutkan dengan Mars Universitas Jambi untuk memperkuat suasana khidmat kebangsaan dan institusional. Setelah sambutan pengantar dan pembacaan doa penutup, inti dari acara ini adalah pengumuman hasil kompetisi yang sangat dinantikan. Pengumuman dilakukan secara bertahap, dimulai dengan penghargaan untuk kategori teknis, yaitu Pemenang Poster Terbaik dan Pemenang Presentasi Terbaik (*Best Presentation*). Puncak acara pengumuman adalah pembacaan daftar pemenang untuk nominasi utama: Juara 1, Juara 2, dan Juara 3 LKTIN CARBON 7.0, yang menjadi penentu capaian seluruh tim finalis. Setelah momen yang menegangkan dan penuh kegembiraan tersebut, rangkaian acara dilanjutkan dengan sesi inspiratif berupa Talkshow yang menghadirkan Kevin Bong, pemenang dari kompetisi bergengsi “Clash of Champions”. Sesi *talkshow* ini memberikan wawasan dan motivasi berharga mengenai perjalanan akademik, riset, dan strategi dalam memenangkan kompetisi ilmiah tingkat nasional. Setelah seluruh rangkaian acara penutupan dan *talkshow* selesai, tim kami segera meninggalkan lokasi acara di UNJA untuk kembali ke penginapan. Waktu yang tersisa digunakan untuk berkemas dan melakukan persiapan yang diperlukan sebelum melanjutkan perjalanan pulang kembali ke Jakarta, mengakhiri seluruh partisipasi dalam Final LKTIN CARBON 7.0 dengan membawa pulang pengalaman berharga dan jaringan baru.

3.2 Kendala yang Ditemukan

Bagian ini menguraikan berbagai kendala yang dihadapi oleh penulis dan tim selama proses pelaksanaan program PRO-STEP, mulai dari tahap awal persiapan proposal hingga penyelesaian seluruh rangkaian kegiatan teknis dan administratif dalam kompetisi LKTIN CARBON 7.0. Selama perjalanan proyek ini, tim menemui sejumlah tantangan yang muncul secara bertahap seiring dengan meningkatnya kompleksitas tugas, beban akademik, serta tuntutan kompetisi. Kendala-kendala tersebut tidak hanya berkaitan dengan aspek teknis seperti pengolahan data dan implementasi algoritma, tetapi juga mencakup tantangan manajerial, koordinasi tim, serta penyesuaian diri terhadap tekanan waktu dan kualitas output yang harus dipenuhi. Dengan demikian, pemaparan kendala ini memberikan gambaran yang lebih realistis mengenai dinamika yang terjadi selama pengerjaan proyek, sekaligus menjadi evaluasi penting untuk meningkatkan efektivitas kerja tim di masa mendatang.

Beberapa tantangan yang paling terasa selama proses pengerjaan antara lain:

1) Menyelaraskan Jadwal Antar Anggota Tim

Salah satu kesulitan paling nyata adalah menemukan waktu yang cocok untuk berdiskusi atau rapat tim. Perbedaan jadwal kuliah, mata kuliah praktikum, dan aktivitas organisasi membuat koordinasi menjadi tidak selalu mudah. Hal ini sering menyebabkan keterlambatan pengerjaan tugas yang membutuhkan kolaborasi langsung.

2) Tingkat Pemahaman yang Berbeda-Beda

Dalam pengerjaan model Machine Learning dan penyusunan laporan ilmiah, tiap anggota memiliki tingkat pemahaman dan kekuatan yang berbeda. Terkadang, sebagian anggota membutuhkan waktu lebih lama untuk memahami konsep tertentu, seperti hyperparameter tuning, analisis hasil, atau

teknik penulisan ilmiah. Hal ini menuntut kami untuk lebih sabar, saling membantu, dan memastikan bahwa setiap bagian dapat terselesaikan dengan standar yang sama.

3) Pengumpulan Referensi yang Lama dan Melelahkan

Mencari jurnal yang benar-benar relevan, kredibel, dan terbaru ternyata menjadi proses yang cukup menyita waktu. Banyak jurnal yang ditemukan tidak sesuai konteks, terlalu teknis, atau tidak mendukung arah penelitian kami. Proses memilah literatur sering kali lebih lama dibandingkan proses coding itu sendiri.

3.3 Solusi atas Kendala yang Ditemukan

Untuk mengatasi berbagai masalah yang muncul selama proses pengerjaan PRO-STEP dan kompetisi LKTIN CARBON 7.0, tim kami menerapkan beberapa langkah strategis yang disesuaikan dengan kondisi serta kemampuan masing-masing anggota. Meskipun tantangan yang muncul beragam, pendekatan solusi yang kami lakukan bertujuan menjaga efisiensi kerja, memperbaiki koordinasi, serta memastikan hasil akhir tetap berkualitas meskipun dikerjakan dalam situasi penuh tekanan.

1) Solusi untuk Menyelaraskan Jadwal Antar Anggota Tim

Untuk mengatasi sulitnya mencocokkan jadwal, kami membuat jadwal kerja fleksibel yang tidak harus mengharuskan seluruh anggota bertemu pada waktu yang sama. Kami memanfaatkan grup chat dan tools kolaborasi seperti Google Docs dan Google Sheets agar setiap anggota bisa mengerjakan bagiannya masing-masing tanpa harus menunggu anggota lain tersedia. Selain itu, kami menetapkan *check-in* harian berdurasi 5–10 menit, entah melalui chat atau call singkat, untuk memastikan progres tetap berjalan. Cara

ini terbukti mengurangi keterlambatan pengerjaan dan membuat koordinasi lebih efisien tanpa perlu rapat yang panjang.

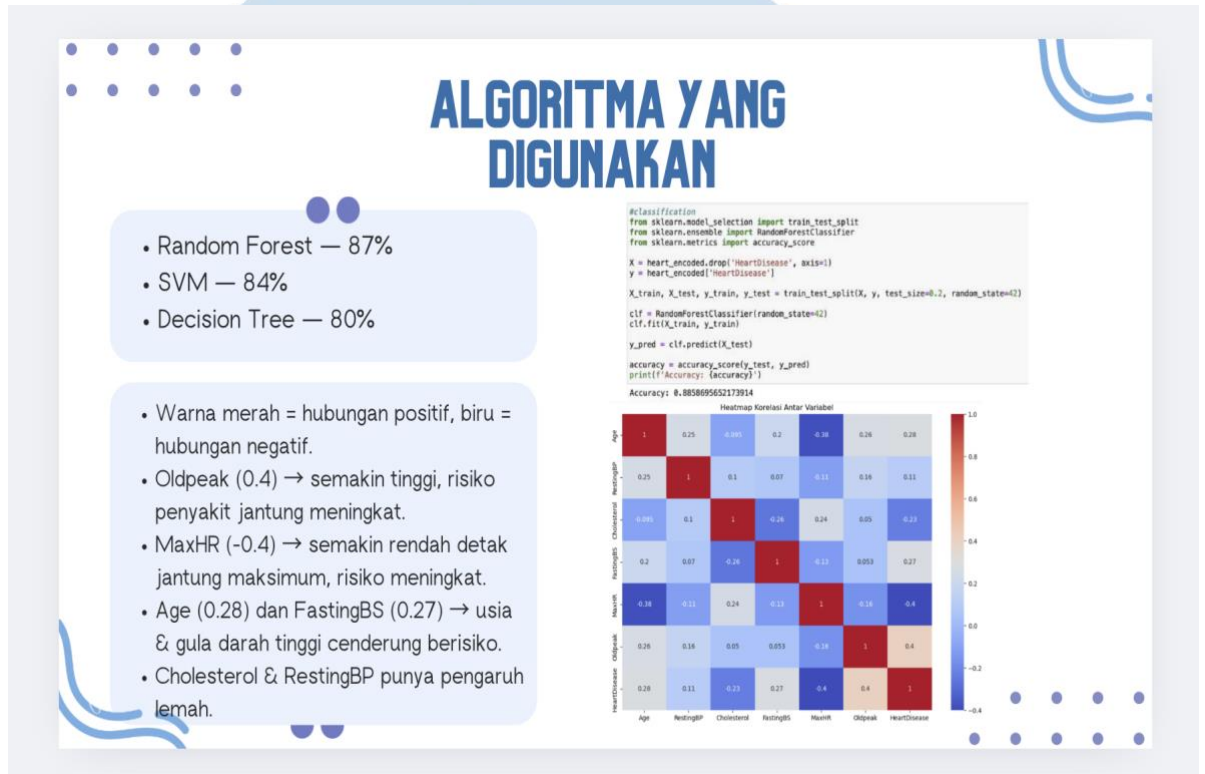
2) Solusi untuk Perbedaan Tingkat Pemahaman Antar Anggota

Untuk menyelesaikan perbedaan pemahaman dalam coding dan penulisan ilmiah, kami menerapkan pendekatan mentoring internal. Anggota yang lebih paham Machine Learning membantu menjelaskan konsep-konsep sulit seperti hyperparameter tuning atau evaluasi model kepada anggota lain melalui diskusi kecil. Sebaliknya, anggota yang lebih kuat dalam penulisan membantu menyempurnakan struktur laporan, tata bahasa, dan penggunaan referensi. Dengan saling mengisi kelebihan masing-masing, proses pengerjaan menjadi lebih seimbang dan setiap anggota dapat berkembang pada bidang yang sebelumnya kurang dikuasai.

3) Solusi untuk Kesulitan dalam Pengumpulan Referensi

Karena mencari jurnal berkualitas membutuhkan waktu lama, kami membagi tugas pencarian referensi menjadi beberapa kategori: jurnal tentang algoritma SVM, jurnal tentang penyakit jantung, jurnal terkait dataset, dan jurnal mengenai metode penelitian. Setiap anggota bertanggung jawab pada kategorinya masing-masing untuk mempercepat proses. Kami juga membuat daftar jurnal dalam spreadsheet yang berisi ringkasan singkat (judul, tahun, metode, dan alasan relevansi) sehingga memudahkan dalam memilih referensi yang akan dimasukkan ke laporan. Pendekatan ini menghemat banyak waktu dan mencegah pencarian jurnal yang berulang.

3.4 Hasil Lomba/Kompetisi



Gambar 3. 9 Contoh Slide PPT Kelompok

Gambar 3.9 Menunjukkan tahap presentasi lomba, tim menampilkan hasil analisis model machine learning yang digunakan dalam proses prediksi penyakit jantung. Slide ini berfokus pada perbandingan performa tiga algoritma utama, yaitu Random Forest, Support Vector Machine (SVM), dan Decision Tree. Dari hasil pengujian, algoritma Random Forest memperoleh akurasi tertinggi sebesar 87%, diikuti oleh SVM dengan 84%, dan Decision Tree dengan 80%. Perbandingan ini menjadi dasar dalam menentukan model yang paling optimal untuk digunakan pada penelitian, yakni Random Forest, meskipun SVM juga menunjukkan performa yang kompetitif.

Pada slide tersebut, ditampilkan pula visualisasi heatmap korelasi antar variabel yang menunjukkan hubungan antar fitur medis dalam dataset. Warna merah pada heatmap menandakan korelasi positif, sedangkan warna biru menunjukkan korelasi negatif.

Beberapa variabel yang diperoleh memiliki pengaruh yang cukup signifikan terhadap risiko penyakit jantung. Misalnya, Oldpeak menunjukkan korelasi positif sebesar 0.4, yang berarti semakin tinggi nilai Oldpeak, semakin besar kemungkinan seseorang mengalami penyakit jantung. Sebaliknya, variabel MaxHR memiliki korelasi negatif (-0.4), yang mengindikasikan bahwa semakin rendah detak jantung maksimum seseorang, semakin tinggi risiko mengalami kondisi tersebut.

Selain itu, fitur Age (0.28) dan FastingBS (0.27) juga menunjukkan hubungan yang cukup relevan terhadap kejadian penyakit jantung, sehingga faktor usia dan kadar gula darah menjadi salah satu indikator penting dalam proses prediksi. Sementara itu, variabel seperti Cholesterol dan RestingBP memiliki korelasi yang lebih rendah sehingga pengaruhnya tidak sebesar variabel lain. Slide ini turut menampilkan cuplikan kode Python yang digunakan dalam pelatihan model Random Forest sebagai bentuk transparansi proses teknis yang dijalankan tim selama kompetisi.

Secara keseluruhan, slide ini menggambarkan proses analisis data, pemilihan model, serta interpretasi hasil yang menjadi bagian penting dalam presentasi final di hadapan dewan juri. Penjelasan ini menegaskan bahwa penelitian yang dilakukan tidak hanya berfokus pada pemodelan prediktif, tetapi juga memahami faktor-faktor medis yang relevan sehingga hasil penelitian dapat memberikan kontribusi yang bermakna dalam konteks deteksi dini penyakit jantung.





Gambar 3. 10 Certificate Top 10 Finalist

Gambar 3.10 Menunjukkan hasil dari rangkaian proses seleksi kompetisi CARBON 7.0, tim yang beranggotakan Marcellinus Gillbert dan Michael Jovando berhasil meraih pencapaian sebagai Top 10 Finalist dalam kategori *Lomba Karya Tulis Ilmiah Nasional (LKTIN)*. Bukti pencapaian ini ditunjukkan melalui sertifikat resmi yang diterbitkan oleh Fakultas Sains dan Teknologi (FST) Universitas Jambi selaku penyelenggara. Sertifikat tersebut diberikan kepada kedua anggota tim sebagai bentuk apresiasi atas

kualitas karya ilmiah yang telah dihasilkan serta keberhasilan tim dalam melewati seluruh tahapan seleksi, mulai dari pengumpulan abstrak, pengiriman full paper, hingga presentasi final di hadapan dewan juri.

Pada sertifikat tersebut tercantum identitas peserta, nomor sertifikat resmi, serta pernyataan bahwa kedua anggota tim berhasil masuk dalam jajaran sepuluh finalis terbaik dari seluruh peserta yang mengikuti kompetisi nasional tersebut. Sertifikat juga ditandatangani langsung oleh pejabat akademik terkait, yaitu Wakil Dekan Bidang Kemahasiswaan dan Alumni FST Universitas Jambi, Ketua Jurusan MIPA FST UNJA, Koordinator Program Studi Kimia FST UNJA, dan Ketua Himpunan Mahasiswa Kimia (HiMKi) UNJA. Tanda tangan para pejabat akademik ini menegaskan validitas serta kredibilitas penghargaan yang diterima oleh tim.

Pencapaian sebagai finalis dalam kompetisi tingkat nasional ini menunjukkan bahwa karya tulis ilmiah yang diajukan—yang berfokus pada pengembangan model prediksi penyakit jantung berbasis teknologi machine learning—memiliki kualitas, relevansi, serta kontribusi ilmiah yang diakui oleh para ahli di bidang sains dan teknologi. Hasil ini tidak hanya memperkuat rekam jejak akademik tim, tetapi juga menjadi pengalaman berharga dalam meningkatkan kemampuan penelitian, penulisan ilmiah, serta presentasi di forum nasional.





Gambar 3. 11 Foto setelah Hasil Pengumuman Lomba

Gambar 3.11 ini menunjukkan dokumentasi saat penulis dan rekan tim, Michael Jovando, menerima sertifikat sebagai Top 10 Finalist pada Lomba Karya Tulis Ilmiah Nasional CARBON 7.0 yang diselenggarakan oleh Universitas Jambi. Pengambilan foto dilakukan di depan area aula utama Universitas Jambi, tepat setelah sesi seminar

nasional dan talkshow yang menjadi rangkaian kegiatan final kompetisi. Pada latar belakang terlihat spanduk resmi acara yang menampilkan tema utama seminar, *“Chemistry for Change: Empowering Youth to Achieve the SDGs”*, beserta para pembicara nasional yang hadir pada kegiatan tersebut.

Foto ini tidak hanya menjadi bukti visual atas pencapaian tim dalam kompetisi tingkat nasional, tetapi juga menggambarkan suasana kegiatan serta momen penghargaan yang diberikan secara langsung oleh panitia. Kehadiran berbagai peserta dari berbagai universitas serta atmosfer akademis yang ditampilkan dalam acara tersebut memberikan pengalaman yang berharga bagi penulis dan tim dalam memperluas wawasan, jejaring, serta kemampuan presentasi ilmiah. Dokumentasi ini sekaligus menegaskan bahwa tim benar-benar mengikuti seluruh rangkaian kegiatan hingga tahap penerimaan penghargaan secara resmi, sehingga memperkuat autentisitas proses keikutsertaan dalam kompetisi CARBON 7.0.



CHEMISTRY FAIR BORN FOR NATIONALISM 7.0

**Pengembangan Sistem Prediksi Penyakit Jantung
Menggunakan Algoritma Machine Learning**



Disusun oleh:

Marcellinus Gillbert	00000068802	2022
Michael Jovando	00000068313	2022

**Universitas Multimedia
Nusantara
Tangerang Selatan
2025**

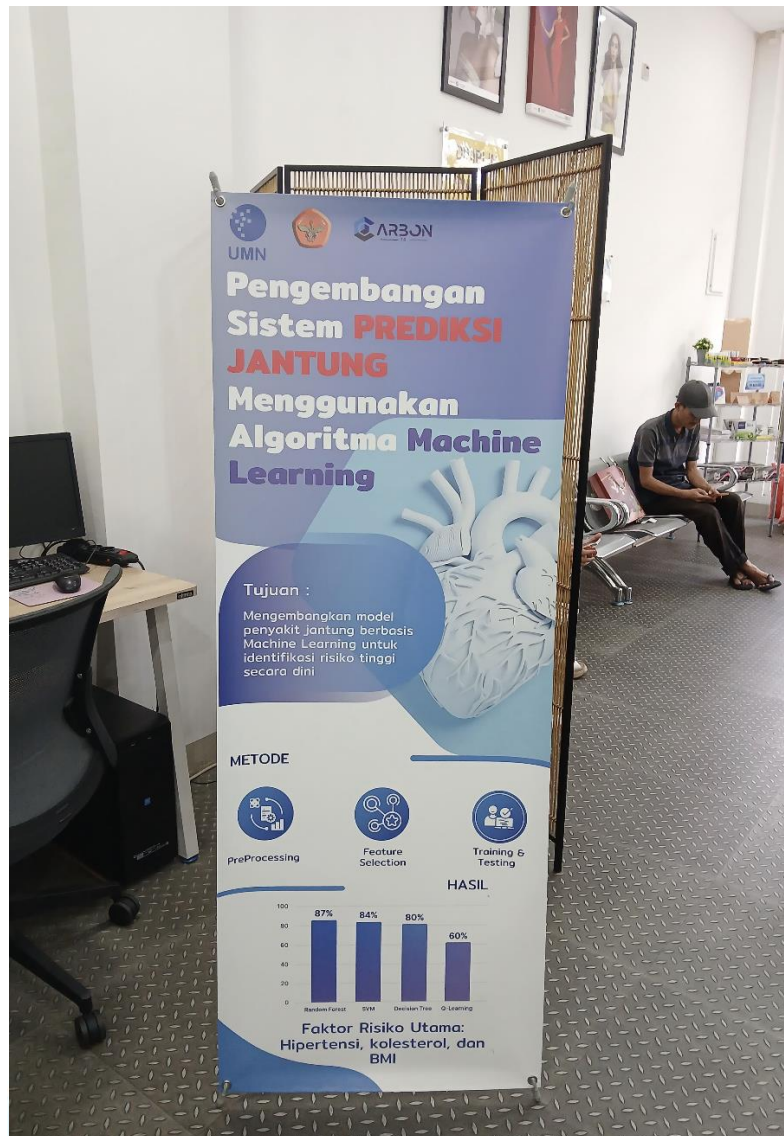
Gambar 3.12 Cover halaman Laporan Lomba

Gambar 3.12 menunjukkan Laporan ilmiah yang disusun sebagai output utama dalam kompetisi Lomba Karya Tulis Ilmiah Nasional (LKTIN) CARBON 7.0 berfungsi sebagai dokumen akademik yang memuat keseluruhan proses penelitian mulai dari

tahap perumusan masalah, metodologi, implementasi model, hingga analisis hasil. Laporan ini terdiri dari beberapa bab yang tersusun secara sistematis sesuai format kompetisi, dengan tujuan untuk memberikan gambaran yang komprehensif mengenai proses PEMODELAN PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE.

Pada bagian sampul laporan Gambar 3.12, tercantum judul penelitian “PEMODELAN PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE”, diikuti dengan identitas penulis beserta logo institusi Universitas Multimedia Nusantara (UMN) sebagai representasi institusi asal peserta. Cover ini merupakan bagian resmi dari dokumen yang dikirimkan kepada panitia CARBON 7.0 sebagai bentuk kelengkapan administratif dan akademik.





Gambar 3. 13 Hasil Poster

Selain penyusunan karya tulis ilmiah (full paper) dan pembuatan model Machine Learning, salah satu output penting yang dihasilkan dalam kompetisi LKTIN CARBON 7.0 adalah standing poster seperti pada gambar 3.13. yang digunakan sebagai media presentasi visual selama sesi penilaian. Standing poster ini dirancang untuk memberikan gambaran komprehensif dan mudah dipahami mengenai keseluruhan penelitian berjudul “PEMODELAN PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE”.

Poster tersebut memuat elemen-elemen inti dari penelitian, dimulai dari judul, tujuan penelitian, hingga penjelasan metode dan hasil utama. Pada bagian metode, poster menampilkan tahapan inti seperti *pre-processing*, *feature selection*, hingga *training & testing* model. Komponen metode disajikan dalam bentuk ikon agar lebih komunikatif dan mudah dipahami oleh juri maupun peserta lain. Selain itu, poster juga menampilkan hasil perbandingan akurasi dari beberapa algoritma (Random Forest, Support Vector Machine, Decision Tree, dan Q-Learning), yang divisualisasikan dalam bentuk grafik batang untuk menunjukkan performa model secara lebih jelas.

Visualisasi tambahan mengenai faktor risiko utama seperti hipertensi, kolesterol, dan BMI turut disertakan untuk memperkuat konteks penelitian dan relevansinya terhadap isu kesehatan masyarakat. Secara keseluruhan, standing poster ini berfungsi sebagai representasi visual dari seluruh proses riset yang dilakukan oleh tim, sekaligus menjadi alat bantu komunikasi yang efektif dalam sesi presentasi dan penjurian di Universitas Jambi.

