

DAFTAR ISI

HALAMAN PENGESAHAN PEMBIMBING	iii
HALAMAN PENGESAHAN PENGUJI.....	iv
SURAT PERNYATAAN KARYA ASLI TUGAS AKHIR.....	v
SURAT PERNYATAAN BEBAS PLAGIASI.....	vi
ABSTRAK	vii
<i>ABSTRACT</i>	viii
PRAKATA	ix
DAFTAR ISI.....	x
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR MODUL PROGRAM	xiv
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah.....	3
1.3. Batasan Masalah	3
1.4. Tujuan Penelitian	3
1.5. Manfaat Penelitian	3
1.6. Tahapan Penelitian	3
1.7. Sistematika Penulisan	4
BAB II TINJAUAN LITERATUR.....	6
2.1. Kualitas Air Sungai	6
2.2. <i>Machine Learning</i>	7
2.2.1 Algoritma Random Forest	8
2.3. <i>Framework</i> CRISP-DM.....	9
2.4. Evaluasi Model	11
2.5. <i>State of The Art</i> (SOTA).....	13
BAB III METODOLOGI PENELITIAN	20
3.1. Metodologi Penelitian.....	20
3.1.1 <i>Business Understanding</i>	21
3.1.2 <i>Data Understanding</i>	22
3.1.3 <i>Data Preparation</i>	23
3.1.4 <i>Modeling</i>	30
3.1.5 <i>Evaluation</i>	31
3.1.6 <i>Deployment</i>	32
BAB IV HASIL DAN PEMBAHASAN	35
4.1. Hasil	35
4.1.1 <i>Business Understanding</i>	36
4.1.2 <i>Data Understanding</i>	37
4.1.3 <i>Data Preparation</i>	43

4.1.4	<i>Modeling</i>	55
4.1.5	<i>Evaluation</i>	55
4.1.6	<i>Deployment</i>	58
4.2.	Pembahasan	62
BAB IV KESIMPULAN DAN SARAN.....		64
5.1.	Kesimpulan	64
5.2.	Saran	64
DAFTAR PUSTAKA		65
LAMPIRAN		67

DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i> pada <i>Multi-class Label</i> (Qorry Meidianingsih et al., 2023) ...	11
Tabel 2.2 Contoh Pengolahan <i>Confusion Matrix</i> untuk Kelas 1 (Qorry Meidianingsih et al., 2023).....	12
Tabel 2.3 <i>State of The Art</i>	13
Tabel 3.1 Daftar Pertanyaan	21
Tabel 3.2 <i>Preview Dataset</i>	21
Tabel 3.3 <i>Preview Dataset</i> (Lanjutan)	22
Tabel 3.4 Deskripsi Parameter <i>Dataset</i> Kualitas Air	22
Tabel 3.5 Deskripsi Parameter <i>Dataset</i> Kualitas Air (Lanjutan)	23
Tabel 3.6 Tabel Skenario <i>Testing</i>	33
Tabel 4.1 <i>Confusion Matrix</i> Hasil <i>Modeling</i>	56
Tabel 4.2 Hasil <i>Feature Importance</i>	56
Tabel 4.3 <i>Confusion Matrix</i> Hasil <i>Modeling</i> Kedua.....	58
Tabel 4.4 Tabel Hasil Pengujian	61

DAFTAR GAMBAR

Gambar 2.1 Jenis dan Algoritma <i>Machine Learning</i>	8
Gambar 2.2 Cara Kerja Algoritma Random Forest (Islam et al., 2021).....	9
Gambar 2.3 <i>Framework</i> CRISP-DM.....	10
Gambar 3.1 Tahapan Penelitian.....	20
Gambar 3.2 <i>Flowchart Data Preparation</i>	24
Gambar 3.3 <i>Flowchart</i> Skenario <i>Preprocessing</i> 1.....	25
Gambar 3.4 <i>Flowchart</i> Skenario <i>Preprocessing</i> 2.....	26
Gambar 3.5 <i>Flowchart</i> Skenario <i>Preprocessing</i> 3.....	27
Gambar 3.6 <i>Flowchart</i> Skenario <i>Preprocessing</i> 4.....	28
Gambar 3.7 <i>Flowchart</i> Penentuan Fitur dan Label	29
Gambar 3.8 <i>Flowchart Data Splitting</i>	30
Gambar 3.9 <i>Flowchart Modeling</i>	31
Gambar 3.10 <i>Wireframe</i> Tampilan Sistem Prediksi Kualitas Air Sungai Kota Yogyakarta	32
Gambar 3.11 <i>Wireframe</i> Halaman Hasil Prediksi.....	33
Gambar 4.1 Tampilan Utama Sistem Prediksi Kualitas Air Sungai Kota Yogyakarta	36
Gambar 4.2 Pratinjau Data	37
Gambar 4.3 Informasi Data	38
Gambar 4.4 Deskripsi Statistik Data	38
Gambar 4.5 <i>Boxplot</i> TSS, BOD, Fosfat, dan pH.....	39
Gambar 4.6 <i>Box Plot</i> DO, COD, FecalColi, dan N.....	40
Gambar 4.7 Grafik Distribusi Nilai Setiap Atribut.....	41
Gambar 4.8 <i>Correlation Matrix</i>	42
Gambar 4.9 Distribusi Data berdasarkan Status.....	43
Gambar 4.10 Distribusi Nilai Kosong pada <i>Dataset</i>	44
Gambar 4.11 <i>Density Plot</i> TSS.....	44
Gambar 4.12 <i>Missing Values</i> pada Atribut ‘Fosfat’ dan ‘FecalColi’	45
Gambar 4.13 Hasil Deteksi <i>Outlier</i> menggunakan IQR.....	46
Gambar 4.14 Hasil Deteksi <i>Outlier</i> dengan Metode MAD	48
Gambar 4.15 Hasil Deteksi <i>Outlier</i> berdasarkan Z-Score.....	49
Gambar 4.16 Hasil Penghapusan <i>Outlier</i>	50
Gambar 4.17 <i>Correlation Matrix</i> Skenario <i>Preprocessing</i> 1.....	50
Gambar 4.18 <i>Correlation Matrix</i> Skenario <i>Preprocessing</i> 2.....	51
Gambar 4.19 <i>Correlation Matrix</i> Skenario <i>Preprocessing</i> 3.....	53
Gambar 4.20 <i>Correlation Matrix</i> Skenario <i>Preprocessing</i> 4.....	54
Gambar 4.21 Halaman Utama Sistem	59
Gambar 4.23 Halaman Hasil Prediksi	59

DAFTAR MODUL PROGRAM

Modul Program 4.1 Proses <i>Import Library</i>	35
Modul Program 4.2 Proses <i>Data Understanding</i>	37
Modul Program 4.3 Proses Imputasi Skenario 1	45
Modul Program 4.4 Proses Deteksi <i>Outlier</i> dengan Metode IQR.....	46
Modul Program 4.5 Proses Deteksi <i>Outlier</i> dengan Metode MAD	47
Modul Program 4.6 Proses Deteksi <i>Outlier</i> berdasarkan Z-Score	48
Modul Program 4.7 Proses Hapus <i>Outlier</i> berdasarkan Z-Score	49
Modul Program 4.8 Proses Imputasi Skenario 2	51
Modul Program 4.9 Proses Imputasi Skenario 3	52
Modul Program 4.10 Proses Penghapusan <i>Outlier</i> Skenario 3	53
Modul Program 4.11 Proses Imputasi Skenario 4	54
Modul Program 4.12 Proses Penentuan Fitur dan Label	55
Modul Program 4.13 Proses <i>Data Splitting</i>	55
Modul Program 4.14 Proses <i>Modeling</i>	55
Modul Program 4.15 Proses <i>Feature Importance</i>	56
Modul Program 4.16 Proses <i>Feature Selection</i>	57
Modul Program 4.17 Proses <i>Deployment</i>	61

DAFTAR LAMPIRAN

LAMPIRAN A. Surat Permohonan Izin Pengambilan Data	67
LAMPIRAN B. Notulensi Hasil Wawancara	68
LAMPIRAN C. <i>Dataset</i> Kualitas Air Sungai dengan Label Status Mutu.....	69