

**KLASIFIKASI MUTU AIR SUNGAI INDONESIA
MENGUNAKAN REGRESI LOGISTIK ORDINAL
BERBASIS *ADAPTIVE SYNTHETIC SAMPLING***

Skripsi

**Disusun untuk memenuhi salah satu syarat
memperoleh gelar sarjana Statistika**

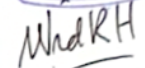
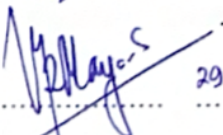


**Erin Naudy Kemalasari
1314622030**

**PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN
ALAM
UNIVERSITAS NEGERI JAKARTA
2026**

**LEMBAR PERSETUJUAN HASIL SIDANG SKRIPSI
KLASIFIKASI MUTU AIR INDONESIA MENGGUNAKAN REGRESI
LOGISTIK ORDINAL BERBASIS *ADAPTIVE SYNTHETIC SAMPLING***

Nama : Erin Naudy Kemalasari
NIM : 1314622030

Nama	Tanda Tangan	Tanggal
Penanggung Jawab Dekan : <u>Dr. Hadi Nasbey, S.Pd., M.Si.</u> NIP. 197909162005011004		4/08/2026
Wakil Penanggung Jawab Wakil Dekan I : <u>Dr. Meiliasari, S.Pd., M.Sc.</u> 197905042009122002		4/08/2026
Ketua Penguji : <u>Dra. Widyanti Rahayu, M.Si.</u> NIP. 196611032001122001		28/07/2026
Sekretaris : <u>Prof. Dr. Ir. Bagus Sumargo, M.Si.</u> NIP. 196309221986011001		20/07/2026
Anggota Pembimbing I : <u>Dr. Vera Maya Santi, S.Si., M.Si.</u> NIP. 197905312005012006		29/07/2026
Pembimbing II : <u>Gorry Meidianingsih, S.Si., M.Si.</u> NIP. 199105192019032019		28/07/2026
Penguji Ahli : <u>Prof. Dr. Dian Handayani, M.Si.</u> NIP. 197404151998032001		22/07/2026

Dinyatakan lulus ujian skripsi tanggal: 16 Juli 2026

LEMBAR PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul “Klasifikasi Status Mutu Air Sungai di Indonesia Menggunakan Regresi Logistik Ordinal Berbasis *Adaptive Synthetic Sampling*” yang disusun sebagai syarat untuk memperoleh gelar Sarjana Statistika dari Program Studi Statistika Universitas Negeri Jakarta adalah karya ilmiah saya dengan arahan dari dosen pembimbing.

Sumber informasi yang diperoleh dari penulis lain yang telah dipublikasikan yang disebutkan dalam teks skripsi ini telah dicantumkan dalam Daftar Pustaka sesuai dengan norma, kaidah, dan etika penulisan ilmiah.

Jika dikemudian hari ditemukan sebagian besar ditemukan sebagian besar skripsi ini bukan hasil karya saya sendiri dalam bagian-bagian tertentu, saya bersedia menerima sanksi pencabutan gelar akademik yang saya sanding dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku.

Jakarta, 14 Juli 2026



A handwritten signature in black ink, identifying the author of the statement.

Erin Naudy Kemalasari



KEMENTERIAN PENDIDIKAN TINGGI, SAINS DAN TEKNOLOGI
UNIVERSITAS NEGERI JAKARTA
PERPUSTAKAAN DAN KEARSIPAN

Jalan Rawamangun Muka Jakarta 13220

Telepon/Faksimili: 021-4894221

Laman: lib.unj.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Erin Naudy Kemalasari
NIM : 1314622030
Fakultas/Prodi : MIPA/Statistika
Alamat email : naudy74@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan dan Kearsipan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

☒ Skripsi ☐ Tesis ☐ Disertasi ☐ Lain-lain (.....)

yang berjudul :

Klasifikasi Status Mutu Air Sungai di Indonesia Menggunakan Regresi Logistik Ordinal

Berbasis Adaptive Synthetic Sampling

Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan dan Kearsipan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 18 Agustus 2026

Penulis

(Erin Naudy Kemalasari)
nama dan tanda tangan

ABSTRAK

ERIN NAUDY KEMALASARI. *Klasifikasi Status Mutu Air Sungai di Indonesia Menggunakan Regresi Logistik Ordinal Berbasis Adaptive Synthetic Sampling.* Skripsi. Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta. Juli 2026.

Mutu air sungai merupakan indikator penting dalam pengelolaan lingkungan. Di Indonesia, status mutu air sungai ditentukan menggunakan Indeks Pencemaran yang mengelompokkan kualitas air menjadi empat kategori, yaitu memenuhi baku mutu, cemaran ringan, cemaran sedang, dan cemaran berat. Ketidakseimbangan jumlah data pada setiap kategori dapat menyebabkan model klasifikasi cenderung memihak kelas mayoritas. Penelitian ini bertujuan menerapkan regresi logistik ordinal untuk mengklasifikasikan status mutu air sungai di Indonesia serta mengevaluasi pengaruh *Adaptive Synthetic Sampling* (ADASYN) dalam meningkatkan performa klasifikasi pada data tidak seimbang dengan membandingkannya terhadap model tanpa penyeimbangan data dan *Synthetic Minority Over-sampling Technique* (SMOTE). Data yang digunakan merupakan data kualitas air sungai dari Kementerian Lingkungan Hidup dan Kehutanan dengan variabel prediktor amonia, *Biological Oxygen Demand* (BOD), *Chemical Oxygen Demand* (COD), *Dissolved Oxygen* (DO), dan Nitrat. Evaluasi model dilakukan menggunakan *confusion matrix* berdasarkan nilai *Accuracy*, *Precision*, *Recall*, dan *F-Measure*. Hasil penelitian menunjukkan bahwa amonia, BOD, COD, dan Nitrat berpengaruh signifikan terhadap status mutu air sungai, sedangkan DO tidak berpengaruh signifikan. Model tanpa penyeimbangan data menghasilkan *Accuracy* tertinggi, sedangkan model berbasis ADASYN menghasilkan *Macro Precision* dan *Macro F-Measure* tertinggi sehingga memberikan performa klasifikasi yang lebih seimbang dibandingkan model tanpa penyeimbangan data maupun SMOTE.

Kata Kunci: Status mutu air sungai, Regresi Logistik Ordinal, *Adaptive Synthetic Sampling*, *Synthetic Minority Over-sampling Technique*, *Confusion Matrix*.

ABSTRACT

ERIN NAUDY KEMALASARI. *Classification of Water Quality Status in Indonesia Using Ordinal Logistic Regression Based on Adaptive Synthetic Sampling. Undergraduate Thesis. Statistics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Negeri Jakarta. July 2026.*

River water quality is a key indicator in environmental management. In Indonesia, water quality status is determined using a Pollution Index that classifies water quality into four categories: meeting quality standards, slightly polluted, moderately polluted, and heavily polluted. An imbalance in the amount of data across each category can cause classification models to favor the majority class. This study aims to apply ordinal logistic regression to classify water quality status in Indonesia and to evaluate the effect of Adaptive Synthetic Sampling (ADASYN) on improving classification performance for imbalanced data by comparing it to a model without data balancing and the Synthetic Minority Over-sampling Technique (SMOTE). The data used consists of river water quality data from the Ministry of Environment and Forestry, with the following predictor variables: ammonia, Biological Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Dissolved Oxygen (DO), and Nitrat. Model evaluation was conducted using a confusion matrix based on Accuracy, Precision, Recall, and F-Measure. The results show that ammonia, BOD, COD, and Nitrat have a significant effect on water quality status, while DO does not have a significant effect. The model without data balancing produced the highest Accuracy, while the ADASYN-based model produced the highest Macro Precision and Macro F-Measure, thus providing more balanced classification performance compared to both the model without data balancing and the SMOTE model.

Keywords: *Water Quality Status, Ordinal Logistic Regression, Adaptive Synthetic Sampling, Synthetic Minority Over-sampling Technique, Confusion Matrix.*

KATA PENGANTAR

Puji dan syukur kehadiran Allah SWT atas segala rahmat, petunjuk, kekuatan, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul " Klasifikasi Status Mutu Air Sungai di Indonesia Menggunakan Regresi Logistik Ordinal Berbasis *Adaptive Synthetic Sampling*. ". Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Statistika pada Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta.

Proses penyusunan skripsi ini tidak terlepas dari berbagai tantangan dan hambatan. Namun, berkat ketekunan, kerja keras, serta dukungan dari berbagai pihak, skripsi ini akhirnya dapat diselesaikan. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Kedua orang tua dan seluruh keluarga yang senantiasa memberikan doa, kasih sayang, dukungan, motivasi, serta semangat kepada penulis.
2. Prof. Dr. Dian Handayani, M.Si. selaku Koordinator Program Studi Statistika FMIPA Universitas Negeri Jakarta yang telah memberikan arahan, motivasi, serta masukan yang sangat berharga selama proses penyusunan skripsi ini.
3. Ibu Dr. Vera Maya Santi, S.Si., M.Si selaku Dosen Pembimbing I yang telah memberikan ilmu, arahan, bimbingan, motivasi, serta masukan yang sangat berharga selama proses penyusunan skripsi ini.
4. Ibu Qorry Meidianingsih, S.Si., M.Si. selaku Dosen Pembimbing II yang telah memberikan ilmu, arahan, bimbingan, motivasi, serta saran yang sangat bermanfaat selama proses penyusunan skripsi ini.
5. Ibu Faroh Ladayya, M.Si. selaku Dosen Pembimbing Akademik yang telah memberikan bimbingan yang sangat berharga selama masa perkuliahan.
6. Penguji Ahli I dan II (*Expert Judgment*) yang telah memberikan masukan, evaluasi, dan penilaian yang sangat berharga dalam proses validasi instrumen penelitian.
7. Seluruh Bapak/Ibu dosen serta staf administrasi Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Jakarta yang telah memberikan ilmu, pelayanan, serta dukungan selama penulis menempuh pendidikan.
8. Kementerian Lingkungan Hidup dan Kehutanan yang telah menyediakan data secara *open source* sehingga penelitian ini dapat terlaksana.
9. Terima kasih kepada pemilik NIM 1314621037 atas segala doa, dukungan, perhatian, motivasi, serta kesabaran yang telah diberikan selama proses penyusunan skripsi ini. Terima kasih telah menjadi tempat berbagi cerita,

memberikan semangat di saat penulis menghadapi berbagai tantangan, serta selalu mendampingi dan meyakinkan penulis untuk tetap berusaha hingga skripsi ini dapat diselesaikan. Segala bentuk dukungan yang diberikan menjadi salah satu penyemangat yang sangat berarti bagi penulis dalam menyelesaikan studi ini.

10. Terima kasih kepada Hani, Jessa, dan Manda selaku sahabat yang telah memberikan dukungan, semangat, doa, serta kebersamaan selama masa perkuliahan hingga proses penyusunan skripsi ini. Terima kasih atas segala bantuan, motivasi, dan momen kebersamaan yang telah diberikan, baik dalam suka maupun duka.
11. Penulis juga mengucapkan terima kasih kepada Alya, Zahra, serta teman-teman dalam grup Tutaya dan teman-teman Program Studi Statistika 2022 atas segala dukungan, doa, semangat, dan kebersamaan yang telah diberikan selama masa perkuliahan hingga penyusunan skripsi ini. Terima kasih atas perhatian, motivasi, bantuan, serta berbagai kenangan berharga yang telah dilalui bersama.
12. Seluruh pihak yang tidak dapat penulis sebutkan satu per satu yang telah memberikan doa, dukungan, bantuan, dan motivasi sehingga skripsi ini dapat diselesaikan.
13. Terakhir, penulis mengucapkan terima kasih kepada diri sendiri atas segala usaha, kerja keras, kesabaran, serta ketekunan dalam menjalani setiap proses penyusunan skripsi ini. Terima kasih karena telah mampu bertahan, bangkit dari berbagai tantangan, dan tidak menyerah hingga akhirnya dapat menyelesaikan salah satu tahap penting dalam perjalanan akademik ini. Semoga segala perjuangan yang telah dilalui menjadi bekal untuk terus berkembang, belajar, dan meraih cita-cita di masa yang akan datang.

Penulis menyadari bahwa skripsi ini masih memiliki berbagai keterbatasan dan jauh dari kata sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun demi penyempurnaan karya ini di masa mendatang. Semoga skripsi ini tidak hanya menjadi salah satu syarat untuk memperoleh gelar sarjana, tetapi juga dapat memberikan manfaat bagi pengembangan ilmu pengetahuan serta menjadi referensi bagi penelitian selanjutnya.

Jakarta, 14 Juli 2026



Erin Naudy Kemalasari

DAFTAR ISI

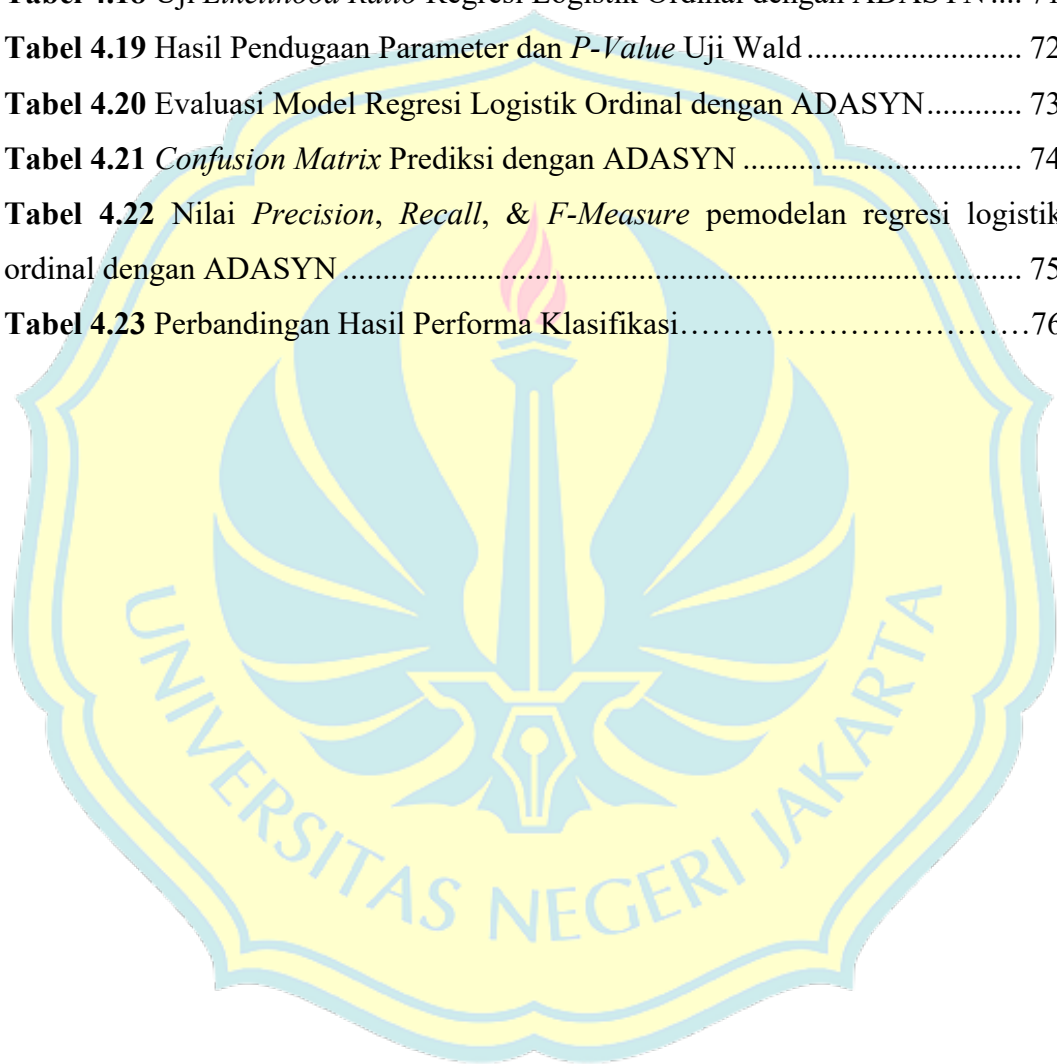
ABSTRAK	i
ABSTRACT	ii
KATA PENGANTAR.....	iii
DAFTAR ISI.....	v
DAFTAR TABEL.....	vii
DAFTAR GAMBAR.....	ix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	7
1.3 Tujuan Penelitian.....	7
1.4 Batasan Masalah.....	8
1.5 Manfaat Penelitian.....	8
BAB II KAJIAN PUSTAKA	9
2.1 <i>Generalized Linear Model</i> (GLM).....	9
2.2 Regresi Logistik Ordinal	15
2.3 Asumsi Regresi Logistik Ordinal.....	19
2.4 Pendugaan Parameter Model Regresi Logistik Ordinal.....	21
2.5 Uji Signifikansi Parameter	23
2.5.1. Uji Rasio <i>Likelihood</i> (<i>Likelihood Ratio Test</i>)	23
2.5.2. Uji Wald	24
2.6 Uji Performa Prediksi Klasifikasi	25
2.7 <i>Synthetic Minority Over-sampling Technique</i> (SMOTE)	28
2.8 <i>Adaptive Synthetic Sampling</i> (ADASYN).....	29
2.9 Mutu air Sungai.....	36
2.9.1. Indikator Kimia	37
BAB III METODOLOGI PENELITIAN	40
3.1 Sumber Data	40
3.2 Prosedur Analisis Data	41
BAB IV HASIL PENELITIAN.....	44
4.1. Eksplorasi Data	44
4.1.1. Sebaran Nitrat berdasarkan Status Mutu Air Sungai	45
4.1.2. Sebaran BOD Status Mutu Air Sungai.....	46
4.1.3. Sebaran COD berdasarkan Status Mutu Air Sungai	47
4.1.4. Sebaran DO berdasarkan Status Mutu Air Sungai	48

4.1.5. Sebaran Amonia berdasarkan Status Mutu Air Sungai.....	49
4.2. Pengecekan multikolinearitas.....	50
4.3. Analisis Regresi Logistik Ordinal.....	50
4.3.1. Penduga Parameter Regresi Logistik Ordinal.....	51
4.3.2. Uji Signifikansi Parameter.....	53
4.3.3. Uji Performa Prediksi Klasifikasi.....	55
4.4. Analisis Regresi Logistik Ordinal dengan SMOTE.....	59
4.4.1. Penduga Parameter Regresi Logistik Ordinal dengan SMOTE.....	60
4.4.2. Uji Signifikansi Parameter.....	62
4.4.3. Uji Performa Prediksi Klasifikasi.....	64
4.5. Analisis Regresi Logistik Ordinal dengan ADASYN.....	68
4.5.1. Penduga Parameter Regresi Logistik Ordinal dengan ADASYN.....	69
4.5.2. Uji Signifikansi Parameter.....	71
4.5.3. Uji Performa Prediksi Klasifikasi.....	73
4.6. Evaluasi Performa Parameter.....	76
BAB V PENUTUP.....	79
5.1. Kesimpulan.....	79
5.2. Saran.....	79
DAFTAR PUSTAKA.....	81
LAMPIRAN.....	86

DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i>	25
Tabel 2.2 Distribusi Kelas Status mutu air sungai Sebelum <i>Oversampling</i>	30
Tabel 2.3 Data Kategori Minoritas Sebelum Proses ADASYN.....	32
Tabel 2.4 Tetangga Terdekat dari Data Minoritas Pertama	32
Tabel 2.5 Perhitungan Jarak Euclidean dan Hasil Pemilihan <i>5-Nearest Neighbor</i> pada Data Minoritas Pertama	33
Tabel 2.6 Nilai Tingkat Kesulitan Data Minoritas dan Jumlah Data Sintetik Metode ADASYN	35
Tabel 2.7 Data Minoritas dan Tetangga Terdekat yang Digunakan dalam Pembentukan Data Sintetik	35
Tabel 2.8 Proses Pembentukan Data Sintetik Menggunakan ADASYN	36
Tabel 2.9 Data Sintetik Hasil <i>Oversampling</i> ADASYN	36
Tabel 3.1 Variabel Penelitian	40
Tabel 3.2 Distribusi Status mutu air sungai.....	41
Tabel 4.1 Hasil Nilai VIF Variabel Prediktor	50
Tabel 4.2 Penduga Parameter dan <i>Odds Ratio</i> Model Regresi Logistik Ordinal ..	51
Tabel 4.3 Estimasi <i>Threshold</i> Model Regresi Logistik Ordinal.....	51
Tabel 4.4 Uji <i>Likelihood Ratio</i> Regresi Logistik Ordinal	54
Tabel 4.5 Hasil Pendugaan Parameter dan <i>P-Value</i> Uji Wald	54
Tabel 4.6 Evaluasi Model Regresi Logistik Ordinal	56
Tabel 4.7 <i>Confusion Matrix</i> Prediksi Tanpa SMOTE dan ADASYN.....	57
Tabel 4.8 Nilai <i>Precision, Recall, & F-Measure</i> pemodelan regresi logistik ordinal	58
Tabel 4.9 Penduga Parameter dan <i>Odds Ratio</i> Model Regresi Logistik Ordinal dengan SMOTE.....	60
Tabel 4.10 Estimasi <i>Threshold</i> Model Regresi Logistik Ordinal dengan SMOTE	60
Tabel 4.11 Uji <i>Likelihood Ratio</i> Regresi Logistik Ordinal dengan SMOTE	63
Tabel 4.12 Hasil Pendugaan Parameter dan <i>P-Value</i> Uji Wald	63
Tabel 4.13 Evaluasi Model Regresi Logistik Ordinal dengan SMOTE.....	65
Tabel 4.14 <i>Confusion Matrix</i> Prediksi dengan SMOTE	66

Tabel 4.15 Nilai <i>Precision</i> , <i>Recall</i> , & <i>F-Measure</i> pemodelan regresi logistik ordinal dengan SMOTE	67
Tabel 4.16 Penduga Parameter dan <i>Odds Ratio</i> Model Regresi Logistik Ordinal dengan ADASYN.....	69
Tabel 4.17 Estimasi <i>Threshold</i> Model Regresi Logistik Ordinal dengan ADASYN	69
Tabel 4.18 Uji <i>Likelihood Ratio</i> Regresi Logistik Ordinal dengan ADASYN	71
Tabel 4.19 Hasil Pendugaan Parameter dan <i>P-Value</i> Uji Wald	72
Tabel 4.20 Evaluasi Model Regresi Logistik Ordinal dengan ADASYN.....	73
Tabel 4.21 <i>Confusion Matrix</i> Prediksi dengan ADASYN	74
Tabel 4.22 Nilai <i>Precision</i> , <i>Recall</i> , & <i>F-Measure</i> pemodelan regresi logistik ordinal dengan ADASYN	75
Tabel 4.23 Perbandingan Hasil Performa Klasifikasi.....	76



DAFTAR GAMBAR

Gambar 1.1 Peta Kualitas Air Sungai berdasarkan IP	2
Gambar 4.1 Distribusi Status Mutu Air Sungai di Indonesia	44
Gambar 4.2 Sebaran Nitrat berdasarkan Status Mutu Air Sungai	45
Gambar 4.3 Sebaran BOD berdasarkan Status Mutu Air Sungai	46
Gambar 4.4 Sebaran COD berdasarkan Status Mutu Air Sungai	47
Gambar 4.5 Sebaran DO berdasarkan Status Mutu Air Sungai	48
Gambar 4.6 Sebaran Amonia berdasarkan Status Mutu Air Sungai	49
Gambar 4.7 Distribusi Status Mutu Air sungai Sebelum & Sesudah Penerapan SMOTE	59
Gambar 4.8 Distribusi Status Mutu Air Sungai Sebelum & Sesudah Penerapan ADASYN	68

