



**IMPLEMENTASI ALGORITMA *EXTRA TREES* UNTUK
KLASIFIKASI CUACA PROVINSI DKI JAKARTA DENGAN
*OVERSAMPLING SMOTE***

SKRIPSI

**RAIHAN KEMMY RACHMANSYAH
NIM. 1910511035**

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA
2023**



**IMPLEMENTASI ALGORITMA *EXTRA TREES* UNTUK
KLASIFIKASI CUACA PROVINSI DKI JAKARTA DENGAN
*OVERSAMPLING SMOTE***

SKRIPSI

**Diajukan sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana
Komputer**

**RAIHAN KEMMY RACHMANSYAH
NIM. 1910511035**

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA
2023**

PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Raihan Kemmy Rachmansyah

NIM : 1910511035

Tanggal : 22 Juni 2023

Judul Skripsi : **Implementasi Algoritma *Extra Trees* untuk Klasifikasi Cuaca
Provinsi DKI Jakarta dengan *Oversampling SMOTE***

Bilamana pada kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 22 Juni 2023

Yang Menyatakan,



Raihan Kemmy Rachmansyah

PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademika Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Raihan Kemmy Rachmansyah
NIM : 1910511035
Fakultas : Ilmu Komputer
Program Studi : Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan karya ilmiah saya kepada Universitas Pembangunan Nasional “Veteran” Jakarta Hak Bebas Royalti Non-Eksklusif (*Non-Exchange Royalty Free Right*) untuk dipublikasikan dengan judul:

IMPLEMENTASI ALGORITMA *EXTRA TREES* UNTUK KLASIFIKASI CUACA PROVINSI DKI JAKARTA DENGAN *OVERSAMPLING SMOTE*

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional “Veteran” Jakarta berhak menyimpan, mengalih media atau memformatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis atau pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 22 Juni 2023

Yang Menyatakan,



Raihan Kemmy Rachmansyah

LEMBAR PENGESAHAN

Tugas Akhir ini diajukan oleh:

Nama : Raihan Kemmy Rachmansyah
NIM : 1910511035
Program Studi : S1 Informatika
Judul Tugas Akhir : Implementasi Algoritma *Extra Trees* untuk Klasifikasi Cuaca Provinsi DKI Jakarta dengan *Oversampling* SMOTE

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Yuni Widiastiwi, S.Kom, M.Si.

Penguji I



Theresia Wati, S.Kom., MTI.

Penguji II



Ria Astriratma, S.Komp., M.Cs.

Pembimbing



Dr. Ermatita, M.Kom.

Dekan



Dr. Widya Cholil, M.T.

Kepala Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : Kamis, 22 Juni 2023



IMPLEMENTASI ALGORITMA *EXTRA TREES* UNTUK KLASIFIKASI CUACA PROVINSI DKI JAKARTA DENGAN *OVERSAMPLING SMOTE*

RAIHAN KEMMY RACHMANSYAH

ABSTRAK

Cuaca merupakan keadaan udara pada suatu wilayah tertentu dengan jangka waktu yang terbatas. Cuaca di Indonesia khususnya di Provinsi DKI Jakarta sangat tidak menentu serta sulit untuk diprediksi. Cuaca yang sulit diprediksi membuat banyak aktivitas warga Provinsi DKI Jakarta terganggu sehingga diperlukan sebuah ilmu teknologi yang diimplementasikan untuk mengklasifikasikan sebuah cuaca. Maka dari itu, penelitian ini menerapkan metode *Machine Learning* untuk klasifikasi cuaca Provinsi DKI Jakarta menggunakan algoritma *Extra Trees* dengan metode *oversampling SMOTE*. Data yang digunakan pada penelitian ini adalah data Prakiraan Cuaca Provinsi DKI Jakarta pada tahun 2017 hingga 2018 yang diperoleh dari situs <https://data.jakarta.go.id/>. Data yang diperoleh memiliki distribusi data yang tidak seimbang sehingga perlu diseimbangkan terlebih dahulu menggunakan metode *resampling* data menggunakan *Synthetic Minority Oversampling Technique* atau *SMOTE*. Berdasarkan hasil penelitian, metode *oversampling SMOTE* tidak mempengaruhi hasil evaluasi pada klasifikasi cuaca Provinsi DKI Jakarta menjadi lebih baik. Hasil evaluasi terbaik didapatkan oleh model menggunakan algoritma *Extra Trees* tanpa metode *oversampling SMOTE* pada rasio 80% data latih dan 20% data uji dengan nilai akurasi sebesar 79,8%, *precision* 63,1%, dan *recall* 56,1%.

Kata Kunci: Klasifikasi, *Extra Trees*, *Oversampling SMOTE*, Cuaca Provinsi DKI Jakarta

IMPLEMENTATION OF THE EXTRA TREES ALGORITHM FOR WEATHER CLASSIFICATION OF DKI JAKARTA PROVINCE WITH OVERSAMPLING SMOTE

RAIHAN KEMMY RACHMANSYAH

ABSTRACT

Weather is the state of the air in a certain area for a limited period of time. The weather in Indonesia, especially in DKI Jakarta Province, is very erratic and difficult to predict. Weather that is difficult to predict disrupts many activities of DKI Jakarta Province residents, so a technological science is needed to classify weather. Therefore, this research applies the Machine Learning method for weather classification in DKI Jakarta Province using the Extra Trees algorithm with the SMOTE oversampling method. The data used in this research is DKI Jakarta Province Weather Forecast data from 2017 to 2018, obtained from the site <https://data.jakarta.go.id/>. The data obtained has an unbalanced data distribution, so it needs to be balanced first using the data resampling method using the Synthetic Minority Oversampling Technique, or SMOTE. Based on the research results, the SMOTE oversampling method does not affect the evaluation results on the weather classification of DKI Jakarta Province for the better. The best evaluation results were obtained by the model using the Extra Trees algorithm without the SMOTE oversampling method at a ratio of 80% training data and 20% test data, with an accuracy value of 79,8%, precision of 63,1%, and recall of 56,1%.

Keywords: *Classification, Extra Trees, Oversampling SMOTE, Weather of DKI Jakarta Province*

KATA PENGANTAR

Puji Syukur peneliti kepada Allah Yang Maha Esa karena rahmat dan karunia-Nya peneliti dapat menyelesaikan skripsi dengan baik sehingga dapat menyelesaikan studi tepat waktu. Penulisan skripsi ini bertujuan untuk memenuhi salah satu persyaratan untuk memperoleh gelar Sarjana Komputer, Jurusan Informatika.

Dalam penulisan skripsi ini, peneliti mendapat banyak dukungan serta bantuan dari berbagai pihak, baik berupa materi, spiritual dan informasi. Pada kesempatan kali ini, peneliti ingin mengucapkan terimakasih kepada:

1. Kedua orang tua dan keluarga yang selalu mendoakan serta mendukung peneliti sehingga dapat menyelesaikan skripsi dengan baik.
2. Ibu Dr. Ermatita, M.Kom. sebagai Dekan Fakultas Ilmu Komputer.
3. Ibu Dr. Widya Cholil, M.I.T. sebagai Kepala Prodi Informatika dan pembimbing akademik yang selalu memberikan informasi serta dukungan selama ini.
4. Ibu Ria Astriratma, S.Komp., M.Cs. sebagai dosen pembimbing skripsi yang berjasa dengan memberikan bimbingan hingga terselesaikannya skripsi ini.
5. Ibu Yuni Widiastiwi, S.Kom, M.Si. dan Ibu Theresia Wati, S.Kom., MTI. sebagai dosen penguji pada sidang skripsi yang telah memberikan saran dan masukkan untuk dapat menyempurnakan skripsi ini menjadi lebih baik.
6. Seluruh jajaran Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta yang telah membantu dalam perizinan dan administrasi.
7. Teman-teman mahasiswa Informatika yang berjuang bersama selama perkuliahan, memberikan semangat dan dorongan untuk dapat menyelesaikan kuliah dan skripsi ini.
8. Sahabat-sahabat perjuangan peneliti yaitu Ajeng, Gito, Ready dan Berli yang selalu membantu dan memberikan masukan serta doa kepada peneliti agar terselesaikannya skripsi ini dengan baik.

9. Terima kasih banyak kepada semua pihak atas bantuan, dukungan, semangat dan doa yang tidak dapat peneliti sebutkan satu persatu.

Peneliti menyadari bahwa penyusunan skripsi ini masih jauh dari sempurna. Oleh karena itu, peneliti mengharapkan kritik dan saran yang bersifat membangun untuk kesempurnaan skripsi ini yang dapat dikirimkan pada *email* peneliti yaitu raikemmy@upnvj.ac.id. Akhir kata peneliti berharap semoga skripsi ini dapat bermanfaat bagi semua pihak.

Jakarta, 05 Juni 2023

Peneliti,

Raihan Kemmy Rachmansyah

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	ii
PERNYATAAN ORISINALITAS	iii
PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	iv
LEMBAR PENGESAHAN	v
ABSTRAK	vi
ABSTRACT.....	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR	xiv
DAFTAR TABEL.....	xvii
DAFTAR LAMPIRAN.....	xx
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Ruang Lingkup.....	3
1.6 Luaran yang Diharapkan	4
1.7 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Cuaca 6	
2.2 <i>Machine Learning</i>	6
2.3 <i>Supervised Learning</i>	7
2.3.1 Klasifikasi	8

2.4	<i>Ensemble Learning</i>	8
	2.4.1 <i>Boostrap Aggregating</i>	8
	2.4.2 <i>Decision Tree</i>	10
	2.4.3 <i>Extra Trees</i>	12
2.5	<i>Data Cleaning</i>	17
	2.5.1 <i>Missing Value</i>	18
	2.5.2 <i>Data Duplikat</i>	18
	2.5.3 <i>Memperbaiki Kesalahan Ketik</i>	18
	2.5.4 <i>Memperbaiki Tipe Data</i>	19
2.6	<i>Feature Engineering</i>	19
	2.6.1 <i>Feature Construction</i>	19
	2.6.2 <i>Feature Selection</i>	20
2.7	<i>Dummy Variable</i>	21
2.8	<i>Imbalance Class</i>	21
	2.8.1 <i>Synthetic Minority Oversampling Technique</i>	22
2.9	<i>Confusion Matrix</i>	24
2.10	<i>Penelitian Terdahulu</i>	27
BAB III METODOLOGI PENELITIAN.....		42
3.1	<i>Tahapan Penelitian</i>	42
3.2	<i>Identifikasi Masalah</i>	43
3.3	<i>Studi Literatur</i>	43
3.4	<i>Pengumpulan Data</i>	43
3.5	<i>Import Library Package</i>	44
3.6	<i>Praproses Data</i>	44
	3.6.1 <i>Data Cleaning</i>	44
	3.6.2 <i>Feature Conctruction</i>	45
	3.6.3 <i>Feature Selection</i>	45

3.6.4	Penyesuaian Data Cuaca Provinsi DKI Jakarta	45
3.6.5	<i>Encoding</i>	46
3.7	Pembagian Data	46
3.7.1	Data Latih dan Data Uji	46
3.8	<i>Resampling Data</i>	47
3.9	Pembuatan Model.....	48
3.10	Evaluasi Hasil.....	49
3.11	Perbandingan Evaluasi Hasil.....	49
3.12	Kesimpulan	50
3.13	Alat Bantu Penelitian	50
3.13.1	Perangkat Keras	50
3.13.2	Perangkat Lunak.....	51
3.14	Jadwal Penelitian.....	51
BAB IV HASIL DAN PEMBAHASAN		54
4.1	Pengambilan Data	54
4.2	<i>Import Library Package</i>	57
4.3	Praproses Data.....	58
4.3.1	<i>Missing Value</i>	58
4.3.2	Data Duplikat	59
4.3.3	Memperbaiki Kesalahan Ketik.....	59
4.3.4	<i>Feature Construction</i>	65
4.3.5	Memperbaiki Tipe Data	69
4.3.6	<i>Feature Selection</i>	71
4.3.7	Penyesuaian Data Cuaca Provinsi DKI Jakarta	74
4.3.8	<i>Encoding dengan Membuat Dummy Variable</i>	77
4.4	Pembagian Data Menjadi Fitur dan Target	80
4.5	Pembagian Data Latih dan Data Uji	83

4.6	Resampling Data Menggunakan <i>Oversampling</i> SMOTE.....	87
4.7	Pemodelan	97
4.8	Evaluasi Hasil.....	114
4.8.1	<i>Confusion Matrix</i> Data Uji Tanpa Metode <i>Oversampling</i> SMOTE pada Rasio 80:20	115
4.8.2	<i>Confusion Matrix</i> Data Latih Tanpa Metode <i>Oversampling</i> SMOTE pada Rasio 80:20	118
4.8.3	<i>Confusion Matrix</i> Data Uji Menggunakan Metode <i>Oversampling</i> SMOTE pada Rasio 80:20.....	122
4.8.4	<i>Confusion Matrix</i> Data Latih Menggunakan Metode <i>Oversampling</i> SMOTE pada Rasio 80:20.....	125
4.8.5	Hasil Evaluasi Seluruh Model <i>Machine Learning</i>	129
4.9	Perbandingan Evaluasi Hasil.....	132
4.10	Analisis Klasifikasi Model <i>Machine Learning</i>	133
4.10.1	Hasil Evaluasi Model pada Rasio 90:10	133
4.10.2	Hasil Evaluasi Model pada Rasio 80:20	136
4.10.3	Hasil Evaluasi Model pada Rasio 70:30	139
4.10.4	Hasil Evaluasi Model pada Rasio 60:40	142
BAB V SIMPULAN DAN SARAN.....		146
DAFTAR PUSTAKA		148
RIWAYAT HIDUP.....		152
LAMPIRAN.....		153

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Contoh Metode <i>Bagging</i> pada <i>Ensemble Learning</i> (Browniee, 2021)	10
Gambar 2.2 Algoritma <i>Extra Trees</i> (Mahkya dkk., 2022).....	14
Gambar 2.3 <i>Pseudo Code</i> Fungsi <i>Split_a_node</i> dan <i>Pick_a_random_split</i> pada Algoritma <i>Extra Tree</i> (Geurts dkk., 2006).....	16
Gambar 2.4 <i>Pseudo Code</i> Fungsi <i>Stop_split</i> pada Algoritma <i>Extra Trees</i> (Geurts dkk., 2006)	16
Gambar 3.1 Tahapan Penelitian	42
Gambar 4.1 Kode Program dan Hasil Pengecekan Data Duplikat.....	59
Gambar 4.2 Jenis dan Jumlah Kategori pada Kolom <i>waktu</i>	60
Gambar 4.3 Kode Program untuk Memperbaiki Kesalahan Ketik pada Kolom <i>waktu</i>	61
Gambar 4.4 Jenis dan Jumlah Kategori pada Kolom <i>cuaca</i>	63
Gambar 4.5 Kode Program untuk Memperbaiki Kesalahan Ketik pada Kolom <i>cuaca</i>	63
Gambar 4.6 Hierarki Kategori Cuaca yang Digunakan	64
Gambar 4.7 Kode Program Tahapan <i>Feature Construction</i> pada Kolom <i>kelembaban_persen</i>	66
Gambar 4.8 Kode Program Tahapan <i>Feature Construction</i> pada Kolom <i>suhu_derajat_celcius</i>	66
Gambar 4.9 Kode Program untuk Memperbaiki Tipe Data yang Tidak Sesuai ...	71
Gambar 4.10 Kode Program Tahapan Penyesuaian Data Cuaca Provinsi DKI Jakarta	75
Gambar 4.11 Kode Program Tahapan <i>Encoding</i> dengan Membuat <i>Dummy Variable</i>	77
Gambar 4.12 Kode Program untuk Duplikasi <i>Dataframe</i>	80

Gambar 4.13 Kode Program Tahapan Pembagian Data Menjadi Fitur dan Target	81
Gambar 4.14 Kode Program Tahapan Pembagian Data pada Rasio 80:20.....	84
Gambar 4.15 Persentase Ukuran Data Setiap Kategori Cuaca	87
Gambar 4.16 Kode Program Tahapan Metode <i>Oversampling</i> SMOTE	91
Gambar 4.17 Pohon Keputusan pada Percabangan Pertama	108
Gambar 4.18 Pohon Keputusan pada Percabangan Kedua	111
Gambar 4.19 Kode Program untuk Membuat Visualisasi Pohon Keputusan	113
Gambar 4.20 Kode Program Tahapan Pemodelan Menggunakan Algoritma <i>Extra Trees</i>	113
Gambar 4.21 Visualisasi Hasil Pohon Keputusan Pertama yang Terbentuk Menggunakan Algoritma <i>Extra Trees</i>	114
Gambar 4.22 <i>Confusion Matrix</i> Data Uji Tanpa Metode <i>Oversampling</i> SMOTE pada Rasio 80:20	115
Gambar 4.23 <i>Confusion Matrix</i> Data Latih Tanpa Metode <i>Oversampling</i> SMOTE pada Rasio 80:20	119
Gambar 4.24 <i>Confusion Matrix</i> Data Uji Menggunakan Metode <i>Oversampling</i> SMOTE pada Rasio 80:20	122
Gambar 4.25 <i>Confusion Matrix</i> Data Latih Menggunakan Metode <i>Oversampling</i> SMOTE pada Rasio 80:20	126
Gambar 4.26 Perbandingan Nilai Akurasi Model Rasio 90:10	135
Gambar 4.27 Perbandingan Nilai <i>Precision</i> Model Rasio 90:10.....	135
Gambar 4.28 Perbandingan Nilai <i>Recall</i> Model Rasio 90:10.....	136
Gambar 4.29 Perbandingan Nilai Akurasi Model Rasio 80:20	138
Gambar 4.30 Perbandingan Nilai <i>Precision</i> Model Rasio 80:20.....	138
Gambar 4.31 Perbandingan Nilai <i>Recall</i> Model Rasio 80:20.....	139
Gambar 4.32 Perbandingan Nilai Akurasi Model Rasio 70:30	141
Gambar 4.33 Perbandingan Nilai <i>Precision</i> Model Rasio 70:30.....	141

Gambar 4.34 Perbandingan Nilai <i>Recall</i> Model Rasio 70:30	142
Gambar 4.35 Perbandingan Nilai Akurasi Model Rasio 60:40	144
Gambar 4.36 Perbandingan Nilai <i>Precision</i> Model Rasio 60:40.....	144
Gambar 4.37 Perbandingan Nilai <i>Recall</i> Model Rasio 60:40	145

DAFTAR TABEL

	Halaman
Tabel 2.1 Tabel Derajat Ketidakseimbangan (Google Developers, 2022)	22
Tabel 2.2 <i>Confusion Matrix</i> pada Klasifikasi <i>Biner</i> (Normawati & Prayogi, 2021)	25
Tabel 2.3 <i>Confusion Matrix</i> pada Klasifikasi <i>Multi-Class</i> (Azis dkk., 2020).....	25
Tabel 2.4 Penelitian Terdahulu	28
Tabel 3.1 Skenario Pembuatan Model	49
Tabel 3.2 Rincian Jadwal Penelitian	52
Tabel 4.1 Data Cuaca Provinsi DKI Jakarta (Data Terbuka Pemerintah Provinsi DKI Jakarta).....	55
Tabel 4.2 Keterangan Setiap Kolom pada Data Cuaca Provinsi DKI Jakarta	56
Tabel 4.3 Informasi Setiap Kolom pada Data Cuaca Provinsi DKI Jakarta	56
Tabel 4.4 Informasi Deskripsi pada Kolom <i>waktu</i> , <i>wilayah</i> dan <i>cuaca</i>	57
Tabel 4.5 <i>Library Package</i> yang Digunakan	57
Tabel 4.6 Jumlah dan Persentase <i>Missing Value</i> pada Setiap Kolom.....	58
Tabel 4.7 Jenis dan Jumlah Kategori pada Kolom <i>waktu</i>	60
Tabel 4.8 Jenis dan Jumlah Kategori pada Kolom <i>waktu</i> Setelah Diperbaiki	61
Tabel 4.9 Jenis dan Jumlah Kategori pada Kolom <i>cuaca</i>	62
Tabel 4.10 Jenis dan Jumlah Kategori pada Kolom <i>cuaca</i> Setelah Diperbaiki	64
Tabel 4.11 Hasil Tahapan <i>Featrue Construction</i> pada Kolom <i>kelembaban_persen</i> dan <i>suhu_derajat_celcius</i>	67
Tabel 4.12 Informasi Setiap Kolom pada Data Cuaca Provinsi DKI Jakarta	69
Tabel 4.13 Tipe Data pada Setiap Kolom	69
Tabel 4.14 Tipe Data pada Setiap Kolom Setelah Diperbaiki	70
Tabel 4.15 Fitur Awal pada Data Cuaca Provinsi DKI Jakarta	71
Tabel 4.16 Hasil Tahapan <i>Feature Selection</i> Menggunakan Metode <i>Wrapper</i>	72
Tabel 4.17 Fitur Setelah Tahapan <i>Feature Selection</i>	73

Tabel 4.18 Informasi Jumlah Data Kategori Waktu pada Setiap Kategori Cuaca	74
Tabel 4.19 Jenis dan Jumlah Kategori pada Kolom <i>cuaca</i> Setelah Data Disesuaikan	75
Tabel 4.20 Data Cuaca Provinsi DKI Jakarta Setelah Tahapan Penyesuaian Data	76
Tabel 4.21 Contoh Penerapan Tahapan <i>Encoding</i> dengan Membuat <i>Dummy Variable</i> pada Kolom <i>waktu</i>	78
Tabel 4.22 Hasil Tahapan <i>Encoding</i> dengan Membuat <i>Dummy Variable</i> pada Kolom <i>waktu</i>	79
Tabel 4.23 Data Target pada Variabel <i>y</i>	81
Tabel 4.24 Data Fitur pada Variabel <i>X</i>	82
Tabel 4.25 Pembagian Rasio antara Data Latih dan Data Uji	83
Tabel 4.26 Hasil Tahapan Pembagian Data Latih pada Rasio 80:20.....	85
Tabel 4.27 Hasil Tahapan Pembagian Data Uji pada Rasio 80:20	86
Tabel 4.28 Sampel Data yang Digunakan Sebagai Contoh Perhitungan Tahapan Metode <i>Oversampling</i> SMOTE pada Rasio 80:20	89
Tabel 4.29 Hasil dari Contoh Tahapan <i>Resampling</i> Data Setelah Menggunakan Metode <i>Oversampling</i> SMOTE pada Rasio 80:20	92
Tabel 4.30 Jumlah Data Setiap Kategori Cuaca Setelah Tahapan <i>Resampling</i> Data pada Rasio 80:20	93
Tabel 4.31 Data pada Variabel <i>y_train</i> Setelah Menggunakan Metode <i>Oversampling</i> SMOTE pada Rasio 80:20.....	93
Tabel 4.32 Data pada Variabel <i>X_train</i> Setelah Menggunakan Metode <i>Oversampling</i> SMOTE pada Rasio 80:20.....	95
Tabel 4.33 Skenario Pembuatan Model	97
Tabel 4.34 Sampel Data yang Digunakan Sebagai Contoh pada Cara Kerja Algoritma <i>Extra Trees</i>	99
Tabel 4.35 Contoh Data pada Tiga Fitur yang Dipilih Secara Acak	100
Tabel 4.36 Pengelompokkan Data pada Kolom <i>kelembaban_mean</i>	101

Tabel 4.37 Pengelompokkan Data pada Kolom <i>suhu_mean</i>	102
Tabel 4.38 Pengelompokkan Data pada Kolom <i>waktu_Malam</i>	102
Tabel 4.39 Jumlah Kategori Cuaca pada Kolom <i>kelembaban_mean</i> Di Bawah Nilai Ambang Batas	104
Tabel 4.40 Jumlah Kategori Cuaca pada Kolom <i>kelembaban_mean</i> Di Atas atau Sama dengan Nilai Ambang Batas.....	104
Tabel 4.41 Jumlah Kategori Cuaca pada Kolom <i>suhu_mean</i> Di Bawah Nilai Ambang Batas	105
Tabel 4.42 Jumlah Kategori Cuaca pada Kolom <i>suhu_mean</i> Di Atas atau Sama dengan Nilai Ambang Batas.....	105
Tabel 4.43 Jumlah Kategori Cuaca pada Kolom <i>waktu_Malam</i> Di Bawah Nilai Ambang Batas	106
Tabel 4.44 Jumlah Kategori Cuaca pada Kolom <i>waktu_Malam</i> Di Atas atau Sama dengan Nilai Ambang Batas.....	106
Tabel 4.45 Data pada <i>Node</i> Pertama yang Terbentuk.....	109
Tabel 4.46 Nilai <i>Confusion Matrix</i> Data Uji Tanpa Metode <i>Oversampling</i> SMOTE pada Rasio 80:20	116
Tabel 4.47 Nilai <i>Confusion Matrix</i> Data Latih Tanpa Metode <i>Oversampling</i> SMOTE pada Rasio 80:20	120
Tabel 4.48 Nilai <i>Confusion Matrix</i> Data Uji Menggunakan Metode <i>Oversampling</i> SMOTE pada Rasio 80:20	123
Tabel 4.49 Nilai <i>Confusion Matrix</i> Data Latih Menggunakan Metode <i>Oversampling</i> SMOTE pada 80:20	126
Tabel 4.50 Hasil Evaluasi Model <i>Machine Learning</i> pada Data Latih.....	130
Tabel 4.51 Hasil Evaluasi Model <i>Machine Learning</i> pada Data Uji	130
Tabel 4.52 Hasil Keseluruhan Evaluasi Model <i>Machine Learning</i>	131

DAFTAR LAMPIRAN

	Halaman
LAMPIRAN 1. HASIL TURNITIN	153