



SKRIPSI

**OPTIMASI MODEL KLASIFIKASI PENYAKIT
JANTUNG MENGGUNAKAN *EXTREME
GRADIENT BOOSTING* DENGAN
HYPERPARAMETER TUNING GRIDSEARCHCV
DAN *BALANCING DATA SMOTE-ENN***

MARCHEL ADIAS PRADANA
NPM 21081010084

DOSEN PEMBIMBING

Dr. Basuki Rahmat, S.Si. MT
Achmad Junaidi, S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**



SKRIPSI

**OPTIMASI MODEL KLASIFIKASI PENYAKIT
JANTUNG MENGGUNAKAN *EXTREME
GRADIENT BOOSTING* DENGAN
HYPERPARAMETER TUNING GRIDSEARCHCV
DAN *BALANCING DATA SMOTE-ENN***

MARCHEL ADIAS PRADANA
NPM 21081010084

DOSEN PEMBIMBING
Dr. Basuki Rahmat, S.Si. MT
Achmad Junaidi, S.Kom, M.Kom

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**

Halaman ini sengaja dikosongkan

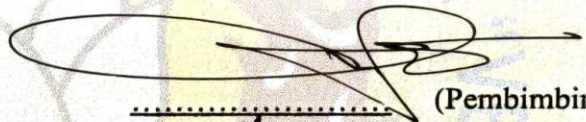
LEMBAR PENGESAHAN

OPTIMASI MODEL KLASIFIKASI PENYAKIT JANTUNG MENGUNAKAN *EXTREME GRADIENT BOOSTING* DENGAN *HYPERPARAMETER TUNING GRIDSEARCHCV* DAN *BALANCING* *DATA SMOTE-ENN*

Oleh :
MARCHEL ADIAS PRADANA
NPM. 21081010084

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 16 Mei 2025

Dr. Basuki Rahmat, S.Si. MT
NIP. 19690723 202121 1 002



(Pembimbing I)

Achmad Junaidi, S.Kom. M.Kom
NPT. 3 7811 04 0199 1



(Pembimbing II)

Budi Nugroho, S.Kom. M.Kom.
NIP. 19800907 202121 1 005



(Ketua Penguji)

Eva Yulia Puspaningrum, S.Kom. M.Kom.
NIP. 19890705 202121 2 002



(Penguji I)

Mengetahui,
Dekan Fakultas Ilmu Komputer



Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

Halaman ini sengaja dikosongkan

LEMBAR PERSETUJUAN

**OPTIMASI MODEL KLASIFIKASI PENYAKIT JANTUNG
MENGUNAKAN *EXTREME GRADIENT BOOSTING* DENGAN
HYPERPARAMETER TUNING GRIDSEARCHCV DAN *BALANCING*
*DATA SMOTE-ENN***

Oleh :

MARCHEL ADIAS PRADANA

NPM. 21081010084

Menyetujui,

Koordinator Program Studi Informatika

Fakultas Ilmu Komputer



Fetty Tri Anggraeny, S.Kom., M.Kom.
NIP. 19820211 202121 2 005

Halaman ini sengaja dikosongkan

SURAT PERNYATAAN BEBAS PLAGIAT

Yang bertandatangan di bawah ini:

Nama Mahasiswa : MARCHEL ADIAS PRADANA
NPM : 21081010084
Program : Sarjana (S1)
Program Studi : Informatika
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabil adikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerimasanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya, 21 Mei 2025
Yang Membuat Pernyataan,




Marchel Adias Pradana
NPM. 21081010084

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa / NPM : MARCHEL ADIAS PRADANA / 21081010084
Judul Skripsi : Optimasi Model Klasifikasi Penyakit Jantung
Menggunakan *Extreme Gradient Boosting*
dengan *Hyperparameter Tuning GridSearchCV*
dan *Balancing Data SMOTE-ENN*
Dosen Pembimbing : 1. Dr. Basuki Rahmat, S.Si. MT
2. Achmad Junaidi, S.Kom, M.Kom

Penyakit jantung menjadi salah satu penyebab utama kematian, khususnya di wilayah Indonesia. Oleh karena itu, dibutuhkan metode klasifikasi untuk meminimalisasi risiko kematian dini akibat penyakit ini. Namun, tantangan muncul karena data kesehatan sering kali bersifat tidak seimbang (*imbalance*), yang dapat memengaruhi performa model klasifikasi. Penelitian ini berfokus pada optimasi model klasifikasi menggunakan algoritma *Extreme Gradient Boosting* (XGBoost) melalui tiga skenario: *baseline* XGBoost, XGBoost dengan *hyperparameter tuning* GridSearchCV, serta XGBoost dengan GridSearchCV yang dikombinasikan dengan teknik *balancing data* SMOTE-ENN.

Model XGBoost dipilih karena kemampuannya dalam membangun model prediktif yang kuat melalui pendekatan *boosting*. GridSearchCV diterapkan untuk mendapatkan kombinasi *hyperparameter* terbaik secara otomatis dengan *cross-validation*. Untuk mengatasi ketidakseimbangan data, teknik SMOTE-ENN digunakan sebagai kombinasi *oversampling* dan *undersampling*, dengan tujuan menyeimbangkan kelas dan mengurangi noise pada data.

Hasil penelitian menunjukkan bahwa *baseline* XGBoost menghasilkan akurasi 0,95 namun *g-mean* rendah pada 0,47. Penerapan GridSearchCV meningkatkan *g-mean* menjadi 0,81 dengan akurasi menurun ke 0,81. Sementara itu, kombinasi GridSearchCV dan SMOTE-ENN menghasilkan *g-mean* 0,79 dengan akurasi lebih stabil di 0,87. Berdasarkan hasil tersebut, model XGBoost dengan GridSearchCV dan SMOTE-ENN dipilih sebagai skenario terbaik karena mampu memberikan keseimbangan antara performa klasifikasi dan akurasi, sekaligus mengatasi ketidakseimbangan data secara efektif. Metode optimasi yang dikembangkan diharapkan dapat menjadi acuan dalam pengembangan model klasifikasi untuk data penyakit, khususnya yang memiliki ketidakseimbangan kelas.

Kata kunci : Penyakit Jantung, Klasifikasi, Ketidakseimbangan Data, XGBoost, GridSearchCV, SMOTE-ENN

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Marchel Adias Pradana / 21081010084
Thesis Title : Optimization of Heart Disease Classification Model
Using Extreme Gradient Boosting with
GridSearchCV Hyperparameter Tuning and
SMOTE-ENN Data Balancing
Advisor : 1. Dr. Basuki Rahmat, S.Si. MT
2. Achmad Junaidi, S.Kom, M.Kom

ABSTRACT

Heart disease is one of the leading causes of death, particularly in Indonesia. Therefore, effective classification methods are needed to minimize the risk of premature death caused by this condition. However, a common challenge in medical data is class imbalance, which can negatively impact the performance of classification models. This study focused on optimizing a classification model using the Extreme Gradient Boosting (XGBoost) algorithm through three scenarios: baseline XGBoost, XGBoost with hyperparameter tuning using GridSearchCV, and XGBoost with GridSearchCV combined with the SMOTE-ENN data balancing technique.

XGBoost was chosen for its capability to build strong predictive models through a boosting approach. GridSearchCV was applied to automatically determine the best hyperparameter combinations using cross-validation. To handle the class imbalance, SMOTE-ENN was used as a hybrid resampling method that combines oversampling of the minority class and also undersampling of the majority class to reduce noise and improve balance.

The results showed that baseline XGBoost achieved an accuracy of 0.95 but had a low geometric mean (g-mean) of 0.47. Incorporating GridSearchCV improved the g-mean to 0.81, although the accuracy dropped to 0.81. Meanwhile, the combination of GridSearchCV and SMOTE-ENN yielded a g-mean of 0.79 with a more stable accuracy of 0.87. Based on these results, XGBoost with GridSearchCV and SMOTE-ENN was selected as the best scenario, offering a balanced performance in terms to increase accuracy and class distribution. The proposed optimization method is expected to become a reference for developing classification models for imbalanced medical datasets.

Keywords: Heart Disease, Classification, Imbalance Data, XGBoost, GridSearchCV, SMOTE-ENN

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, petunjuk, dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul **“Optimasi Model Klasifikasi Penyakit Jantung Menggunakan *Extreme Gradient Boosting* dengan *Hyperparameter Tuning GridSearchCV* dan *Balancing Data SMOTE-ENN*”** dengan baik.

Penyusunan skripsi ini merupakan salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur. Penulis menyampaikan apresiasi yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan dan kontribusi selama proses penelitian ini. Penulis menyadari bahwa penyelesaian skripsi ini tidak lepas dari bantuan, bimbingan, serta dorongan baik secara moril, spiritual, maupun materiil dari berbagai pihak. Oleh karena itu, penulis menyampaikan ucapan terima kasih kepada:

1. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT., selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Fetty Tri Anggraeny, S.Kom, M.Kom, selaku Ketua Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Bapak Dr. Basuki Rahmat, S.Si. MT, selaku Dosen Pembimbing I, yang telah membimbing penulis selama pengerjaan penelitian ini.
4. Bapak Achmad Junaidi, S.Kom, M.Kom, selaku Dosen Pembimbing II, yang telah membimbing penulis selama pengerjaan penelitian ini.
5. Bapak Budi Nugroho, S.Kom., M.Kom., selaku Dosen Penguji I, yang telah menguji penulis dengan memberikan perspektif baru pada penelitian ini.
6. Ibu Eva Yulia Puspaningrum, S.Kom., M.Kom., selaku Dosen Penguji II, yang telah memberikan evaluasi terkait hasil penelitian ini.
7. Bapak Firza Prima Aditiawan, S.Kom., MTI, selaku Dosen Wali, yang telah mengarahkan penulis selama perkuliahan dan penelitian ini.

8. Dosen dan Staff Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur atas segala bantuan untuk mempermudah penulis dalam penyusunan penelitian ini.
9. Ayah, Mama, serta Adik tercinta yang tidak pernah bosan untuk mendoakan dan memberikan semangat, serta dukungan materiil kepada penulis selama menempuh studi.
10. Teman-teman yang telah memberikan dukungan selama keberlangsungan masa perkuliahan dan penyusunan penelitian.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat berbagai kekurangan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari berbagai pihak guna menyempurnakan karya tulis ini. Pada akhirnya, dengan segala keterbatasan yang ada, penulis berharap laporan ini dapat memberikan manfaat, baik bagi pembaca secara umum maupun bagi penulis secara khusus.

Surabaya, 21 Mei 2025

Penulis,



Marchel Adias Pradana
NPM. 21081010084

DAFTAR ISI

LEMBAR PENGESAHAN	i
LEMBAR PERSETUJUAN	iii
SURAT PERNYATAAN ORISINALITAS	iv
ABSTRAK	vii
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xiii
DAFTAR GAMBAR.....	xvii
DAFTAR TABEL	xix
DAFTAR NOTASI.....	xxi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Tujuan Penelitian.....	4
1.4. Batasan Masalah.....	5
1.5. Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA	7
2.1. Penelitian Terdahulu.....	7
2.2. Penyakit Jantung.....	9
2.3. Data Mining.....	10
2.4. Seleksi Fitur Chi-Square	11
2.5. Imbalance Data.....	12
2.5.1 SMOTE (Synthetic Minority Oversampling Technique).....	13
2.5.2. ENN (Edited Nearest Neighbors).....	15
2.5.3. SMOTEENN	17
2.6. XGBoost (Extreme Gradient Boosting)	18
2.7. GridSearchCV	21

2.8. Confusion Matrix	22
BAB III METODOLOGI	25
3.1. Tahapan Penelitian	25
3.2. Studi Literatur.....	26
3.3. Analisis Kebutuhan	26
3.3.1. Spesifikasi Perangkat Keras	26
3.3.2. Spesifikasi Perangkat Lunak	27
3.4. Pengumpulan Data	27
3.5. Preprocessing Data	29
3.6. Pembuatan Model.....	39
3.6.1. Teknik <i>Oversampling</i> (SMOTE).....	40
3.7.2. Teknik Undersampling (ENN).....	41
3.6.3. Teknik <i>Hybridsampling</i> (SMOTEENN)	42
3.7. Model XGBoost dengan Hyperparameter Tuning	43
3.8. Evaluasi Model.....	44
3.9. Skenario Pengujian.....	44
3.10. Analisis Hasil	46
BAB IV PENGUJIAN DAN ANALISA	47
4.1. Preprocessing Data	47
4.1.1. <i>Exploratory Data Analysis</i> (EDA)	47
4.1.2. Dataset Cleaning	54
4.1.3. Feature Selection	57
4.1.4. Transformation	59
4.1.5. Dataset Splitting	64
4.2. Balancing Data	65
4.3. Pembuatan Model.....	67
4.3.1. Model XGBoost	67
4.3.2. Model XGBoost dengan GridSearchCV	68
4.3.3. Model XGBoost dengan GridSearchCV dan SMOTE-ENN	69
4.4. Evaluasi Model.....	69
4.5. Skenario Pengujian.....	75
4.5.1. Perubahan Rasio Data 70:30	75

4.5.2. Perubahan Rasio Data 75:25	76
4.5.3. Perubahan Rasio Data 80:20	78
4.6. Analisis Hasil	79
4.6.1. Perbandingan Waktu Komputasi.....	79
4.6.2. Persebaran Performa Skenario	81
4.6.3. Perbandingan Rata – Rata Tiap Metode.....	82
4.6.4. Perbandingan Skenario Terbaik Tiap Metode.....	84
BAB V.....	87
5.1. Kesimpulan.....	87
5.2. Saran Pengembangan	87
DAFTAR PUSTAKA	89

Halaman ini sengaja dikosongkan

DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi Teknik <i>Oversampling</i>	14
Gambar 2. 2 Ilustrasi SMOTE	14
Gambar 2. 3 Ilustrasi Teknik <i>Undersampling</i>	15
Gambar 2. 4 Ilustrasi ENN.....	16
Gambar 2. 5 Ilustrasi Teknik <i>Hybridsampling</i>	17
Gambar 2. 6 Ilustrasi SMOTEENN	18
Gambar 2. 7 Ilustrasi XGBoost.....	19
Gambar 2. 8 Ilustrasi Grid Search Cross Validation.....	21
Gambar 3. 1 Alur Desain Sistem.....	25
Gambar 3. 2 Flowchart Preprocessing Data.....	30
Gambar 3. 3 Visualisasi Data Penyakit Jantung	32
Gambar 3. 4 Flowchart Pembuatan Model Klasifikasi	39
Gambar 3. 5 Flowchart Teknik SMOTE.....	40
Gambar 3. 6 Flowchart Teknik ENN	41
Gambar 3. 7 Flowchart Teknik SMOTEENN	42
Gambar 4. 1 Struktur dan Isi pada Dataset Penyakit Jantung.....	47
Gambar 4. 2 Distribusi Penyakit Jantung.....	48
Gambar 4. 3 Distribusi Penyakit Jantung Berdasarkan Jenis Kelamin.....	49
Gambar 4. 4 Distribusi Penyakit Jantung Berdasarkan Kelompok Usia	49
Gambar 4. 5 Distribusi Penyakit Jantung Berdasarkan Kondisi Kesehatan	50
Gambar 4. 6 Distribusi Penyakit Jantung Berdasarkan Aktivitas Fisik.....	51
Gambar 4. 7 Distribusi Penyakit Jantung Berdasarkan Status Diabetes.....	51
Gambar 4. 8 Distribusi Penyakit Jantung Berdasarkan Riwayat Stroke.....	52
Gambar 4. 9 Distribusi Penyakit Jantung Berdasarkan Status Merokok	53
Gambar 4. 10 Jumlah Sampel Data Duplikat.....	54
Gambar 4. 11 Jumlah Sampel Data Duplikat Setelah Dilakukan Pembersihan....	55
Gambar 4. 12 Jumlah <i>Missing Value</i>	55
Gambar 4. 13 Sampel Data Sebelum Dilakukan Penanganan <i>Missing Value</i>	56
Gambar 4. 14 Sampel Data Sebelum Dilakukan Penanganan <i>Missing Value</i>	56
Gambar 4. 15 Hasil Setelah Dilakukan Penanganan <i>Missing Value</i>	57
Gambar 4. 16 Fitur Teratas Beserta Nilai <i>Chi-Square</i> dan <i>P-value</i> Tiap Fitur.....	58

Gambar 4. 17 Hasil Akhir <i>Feature Selection Chi-square</i>	59
Gambar 4. 18 Sampel Data Sebelum dan Sesudah <i>Binary Encoding</i>	60
Gambar 4. 19 Sampel Data Sebelum dan Sesudah <i>Ordinal Encoding</i>	62
Gambar 4. 20 Sampel Data Sebelum <i>One-Hot Encoding</i>	63
Gambar 4. 21 Sampel Data Sesudah <i>One-Hot Encoding</i>	63
Gambar 4. 22 Hasil Akhir <i>Transformation</i> Pada Dataset	64
Gambar 4. 23 Hasil pembagian rasio dataset 70:30	64
Gambar 4. 24 Hasil pembagian rasio dataset 75:25	65
Gambar 4. 25 Hasil pembagian rasio dataset 80:20	65
Gambar 4. 26 Contoh Hasil Teknik Kombinasi SMOTE-ENN.....	66
Gambar 4. 27 <i>Hyperparameter Tuning</i> Model XGBoost tanpa <i>Balancing Data</i> .	69
Gambar 4. 28 <i>Hyperparameter Tuning</i> Model XGBoost dengan SMOTE-ENN.	69
Gambar 4. 29 Evaluasi Model XGBoost.....	71
Gambar 4. 30 Confusion Matrix Model XGBoost.....	71
Gambar 4. 31 Evaluasi Model XGBoost dengan GridSearchCV	72
Gambar 4. 32 <i>Confusion Matrix</i> Model XGBoost dengan GridSearchCV	72
Gambar 4. 33 Evaluasi Model XGBoost dengan GridSearchCV dan SMOTE-ENN	73
Gambar 4. 34 Confusion Matrix Model XGBoost dengan GridSearchCV dan SMOTE-ENN.....	74
Gambar 4. 35 Perbandingan Waktu Komputasi Tiap Metode	80
Gambar 4. 36 Persebaran Nilai Akurasi Tiap Skenario Ketiga Metode	81
Gambar 4. 37 Persebaran Nilai G-Mean Tiap Skenario Ketiga Metode.....	82
Gambar 4. 38 Perbandingan Rata-Rata Ketiga Metode	83
Gambar 4. 39 Perbandingan Hasil Skenario Terbaik Ketiga Metode.....	84

DAFTAR TABEL

Tabel 2. 1 Tabel Kontigensi	11
Tabel 2. 2 Contoh Penghapusan ENN.....	17
Tabel 2. 3 <i>Confusion Matrix</i>	22
Tabel 3. 1 Spesifikasi Perangkat Keras.....	26
Tabel 3. 2 Spesifikasi Perangkat Lunak.....	27
Tabel 3. 3 Rincian Atribut pada Dataset	27
Tabel 3. 4 Dataset Pasien	31
Tabel 3. 5 Contoh Analisis Tipe Data	31
Tabel 3. 6 Contoh Pengecekan Missing Value	31
Tabel 3. 7 Distribusi Data Penyakit Jantung.....	31
Tabel 3. 8 Contoh Sebelum Penanganan Duplikasi Data	32
Tabel 3. 9 Contoh Sesudah Penanganan Duplikasi Data	33
Tabel 3. 10 Contoh Sebelum Penghapusan Missing Value	33
Tabel 3. 11 Contoh Sesudah Penghapusan Missing Value	33
Tabel 3. 12 Contoh Sebelum Penggantian Missing Value.....	34
Tabel 3. 13 Contoh Sesudah Penggantian Missing Value	34
Tabel 3. 14 Contoh Sebelum Dilakukan Seleksi Fitur	35
Tabel 3. 15 Contoh Sebelum Dilakukan Seleksi Fitur	35
Tabel 3. 16 Contoh Sebelum <i>Binary Encoding</i>	36
Tabel 3. 17 Contoh Sesudah <i>Binary Encoding</i>	36
Tabel 3. 18 Contoh Sebelum <i>Ordinal Encoding</i>	36
Tabel 3. 19 Contoh Sesudah <i>Ordinal Encoding</i>	37
Tabel 3. 20 Contoh Sebelum <i>One-hot Encoding</i>	37
Tabel 3. 21 Contoh Sesudah <i>One-hot Encoding</i>	37
Tabel 3. 22 Pembagian pada Fitur Dataset.....	38
Tabel 3. 23 Pembagian pada Target Dataset.....	38
Tabel 3. 24 Nilai Hyperparameter yang Diuji.....	43
Tabel 3. 25 Skenario Pengujian	46
Tabel 4. 1 Hasil Evaluasi Rasio Data 70:30.....	75
Tabel 4. 2 Hasil Evaluasi Rasio Data 75:25.....	76
Tabel 4. 3 Hasil Evaluasi Rasio Data 80:20.....	78

Tabel 4. 4 Hasil Perhitungan Waktu Eksekusi.....	79
--	----

DAFTAR NOTASI

X^2	:	Nilai <i>Chi-square</i>
O	:	frekuensi observasi
E	:	frekuensi ekspektasi
T_{bi}	:	total nilai baris ke-i
T_{ki}	:	total nilai kolom ke-i
N	:	total keseluruhan nilai observasi
dk	:	derajat kebebasan
$d(x, y)$	:	jarak data x dengan data y
n	:	jumlah dimensi atribut
x_i	:	atribut ke-i untuk data x
y_i	:	atribut ke-i untuk data y
x_{baru}	:	data sintetik yang dihasilkan
x_{knn}	:	tetangga terdekat dari x_i
δ	:	nilai acak antara 0 dan 1
$\hat{y}_0$	:	prediksi awal untuk semua data
y	:	Nilai target
r_i	:	nilai residual ke-i
y_i	:	nilai target aktual data ke-i
$\sum r$	:	Jumlah residual
$\sum P_{i-1}$	:	jumlah probabilitas sebelumnya
λ	:	parameter regularisasi
$\hat{y}_i^{(t)}$	:	prediksi baru pada iterasi t
$\hat{y}_i^{(t-1)}$	:	prediksi pada iterasi sebelumnya
η	:	learning rate, mengontrol besar perubahan nilai prediksi
$f_t(x_i)$	:	output dari tree untuk data point x_i pada iterasi t

Halaman ini sengaja dikosongkan