



**PENERAPAN SMOTE UNTUK MENANGANI KETIDAKSEIMBANGAN  
KELAS DALAM KLASIFIKASI TINGKAT KEMUNGKINAN OBESITAS  
MENGUNAKAN EXTREME LEARNING MACHINE (ELM)**

**SKRIPSI**

**Untuk Memenuhi Persyaratan Dalam  
Menyelesaikan Sarjana Strata S1-Ilmu Komputer**

**Oleh**

**Titin Supriyatini**

**NIM 1811016120012**

**PROGRAM STUDI S-1 ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMBUNG MANGKURAT  
BANJARBARU  
MARET 2023**

# SKRIPSI

## PENERAPAN SMOTE UNTUK MENANGANI KETIDAKSEIMBANGAN KELAS DALAM KLASIFIKASI TINGKAT KEMUNGKINAN OBESITAS MENGUNAKAN EXTREME LEARNING MACHINE (ELM)

Oleh :

**TITIN SUPRIYATIN**  
**1811016120005**

Telah dipertahankan di depan Dosen Penguji pada tanggal 27 Maret 2023:

Susunan Dosen Penguji :

**Pembimbing I**



Triando H. Saragih, S.Kom., M.Kom

NIP. 199308242019031012

**Dosen Penguji I**



Dwi Kartini, S.Kom., M.Kom

NIP. 198704212012122003

**Pembimbing II**



Andi Farmadi, S.Si., M.T

NIP. 197307252008011006

**Dosen Penguji II**



Rahmat Ramadhani, S.Kom., M.Sc

NIP. 19920330201901110001



Banjarmasin, 27 Maret 2023

**Kena Program Studi Ilmu Komputer**

Iwan Budiman, S.T., M.Kom

NIP. 197703252008121001

## **PERNYATAAN**

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan disuatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, Maret 2023



Titin Supriyatin

NIM. 1811016120012

## ABSTRAK

### **PENERAPAN SMOTE UNTUK MENANGANI KETIDAKSEIMBANGAN KELAS DALAM KLASIFIKASI TINGKAT KEMUNGKINAN OBESITAS MENGGUNAKAN EXTREME LEARNING MACHINE (ELM)**

(Oleh: Titin Supriyatin; Pembimbing: Triando H. Saragih, S.Kom., M.Kom dan Andi Farmadi, S.Si., M.T. ; 2022; 61 halaman)

Berbagai macam penyakit degeneratif seperti penyakit jantung dan stroke, dapat muncul akibat dari obesitas. Oleh karena itu, penting adanya pengetahuan dalam kasus tingkat obesitas sehingga dapat dilakukan penanganan dini, caranya yaitu dengan teknik klasifikasi. ELM merupakan metode pembelajaran baru dari jaringan syaraf tiruan dengan single hidden layer atau biasa disebut *single hidden layer feedforward neural network*. Data obesitas yang digunakan dalam penelitian ini merupakan data tidak seimbang, pada beberapa kasus hal ini dapat berpengaruh terhadap hasil klasifikasi. SMOTE merupakan metode *oversampling* yang sudah terbukti dapat menangani data tidak seimbang. Tujuan penelitian ini adalah untuk mengetahui seberapa besar pengaruh SMOTE dalam klasifikasi tingkat kemungkinan obesitas menggunakan ELM. Dari pegujian yang telah dilakukan didapatkan bahwa kinerja paling tinggi pada data tanpa penambahan SMOTE didapat pada jumlah 100 *hidden neuron* dengan akurasi 67% dan AUC 0,81. Sedangkan untuk data dengan penambahan SMOTE diperoleh hasil paling tinggi pada 500 *hidden neuron* dengan akurasi 84% dan AUC 0,93.

Kata Kunci : *Obesitas, Extreme Learning Machine, SMOTE*

## ABSTRACT

### **IMPLEMENTATION OF SMOTE TO TREAT CLASS IMBALANCE IN OBESITY PROBABILITY CLASSIFICATION USING EXTREME LEARNING MACHINE (ELM)**

(By: Titin Supriyatin; Mentor: Triando H. Saragih, S.Kom., M.Kom and Andi Farmadi, S.Si., M.T. ; 2023; 61 halaman)

*Various kinds of degenerative diseases such as heart disease and stroke, can arise as a result of obesity. Therefore, it is important to have knowledge in cases of obesity so that early treatment can be carried out, the method is by using classification techniques. ELM is a new learning method of artificial neural networks with a single hidden layer or commonly called a single hidden layer feedforward neural network. Obesity data used in this study is unbalanced data, in some cases this can affect the classification results. SMOTE is an oversampling method that has been proven to be able to handle unbalanced data. The purpose of this study was to determine how much influence SMOTE had on the classification of the probability of obesity using ELM. From the tests that have been carried out, it was found that the highest performance on data without the addition of SMOTE was obtained on a total of 100 hidden neurons with an accuracy of 67% and an AUC of 0.81. As for the data with the addition of SMOTE, the highest results were obtained for 500 hidden neurons with an accuracy of 84% and an AUC of 0.93.*

Keywords : *Obesity, Extreme Learning Machine, SMOTE*

## PRAKATA

Assalamualaikum Warahmatullahi Wabarakatuh

Puji syukur penulis panjatkan ke hadirat Allah SWT karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul “Penerapan Smote Untuk Menangani Ketidakseimbangan Kelas Dalam Klasifikasi Tingkat Kemungkinan Obesitas Menggunakan *Extreme Learning Machine* (ELM)” untuk memenuhi syarat dalam menyelesaikan Pendidikan S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat. Tak lupa penulis panjatkan sholawat dan salam ke hadirat Rasulullah Muhammad SAW beserta para sahabat, keluarga, dan pengikut beliau hingga yaumul qiama.

Pada lembar ini penulis ingin menyampaikan ucapan terimakasih kepada pihak- pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan skripsi ini, adapun yang dimaksud adalah sebagai berikut:

1. Keluarga terutama orang tua yaitu Bapak Supriyadi dan Ibu Herna yang selalu memberikan bantuan, semangat, doa, dan dukungan dalam proses menyelesaikan skripsi ini.
2. Bapak Triando H. Saragih, S.Kom., M.Kom selaku dosen pembimbing utama yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
3. Bapak Andi Farmadi, S.Si., M.T selaku dosen pembimbing pendamping yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
4. Bapak Irwan Budiman S.T., M.Kom selaku Ketua Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan izin beliau skripsi ini dapat diselesaikan.
5. Bapak M. Reza Faisal., S.T., M.T., Ph.D selaku dosen pembimbing akademik dan seluruh Dosen dan staf Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.

6. Teman-teman dan kerabat khususnya teman-teman Ilmu Komputer 2018, yang sudah ikut serta dalam memberikan semangat, memberikan doa serta membantu mengerjakan penelitian ini.
7. Semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam penyelesaian skripsi ini.

Penulis berharap skripsi ini dapat bermanfaat memberikan sumbangan pikiran dan ilmu pengetahuan baik bagi penulis sendiri maupun bagi pembaca. Penulis menyadari bahwa skripsi ini jauh dari kata sempurna, oleh karena itu penulis mengharapkan kritik dan saran yang membangun guna kesempurnaan skripsi ini untuk perbaikan dikemudian hari.

Banjarbaru, Maret 2023



Titin Supriyatin

NIM. 1811016120012

## DAFTAR ISI

SKRIPSI.....	ii
PERNYATAAN.....	iii
ABSTRAK .....	iv
ABSTRACT.....	v
PRAKATA.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR .....	xii
DAFTAR LAMPIRAN.....	xiii
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan.....	3
1.5 Manfaat Penelitian.....	4
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Kajian Terdahulu .....	5
2.2 Keaslian Penelitian .....	7
2.3 Landasan Teori .....	8
2.3.1. Obesitas .....	8
2.3.2. Data Mining .....	10
2.3.3. Normalisasi Min-Max .....	11
2.3.4. Cross Validation.....	12
2.3.5. Synthetic Minority Oversampling Technique .....	13
2.3.6. Extreme Learning Machine.....	17
2.3.7. Evaluasi .....	20
<b>BAB III METODE PENELITIAN</b>	
3.1 Alat Penelitian .....	23

3.2	Bahan Penelitian.....	23
3.3	Variabel Penelitian .....	25
3.4	Prosedur Penelitian.....	25
<b>BAB IV HASIL DAN PEMBAHASAN</b>		
4.1	Hasil.....	27
4.1.1.	Pengumpulan Data .....	27
4.1.2.	Preprocessing Data.....	28
4.1.3.	Pembagian Data .....	29
4.1.4.	Klasifikasi .....	31
4.1.5.	Evaluasi .....	41
4.1.6.	Perbandingan Kinerja.....	46
4.2	Pembahasan .....	48
<b>BAB V PENUTUP</b>		
5.1	Kesimpulan.....	52
5.1	Saran.....	52
DAFTAR PUSTAKA .....		54
RIWAYAT HIDUP.....		62

## DAFTAR TABEL

<b>Tabel</b>	<b>Halaman</b>
Tabel 1. Keaslian Penelitian.....	7
Tabel 2. Rancangan Penelitian.....	8
Tabel 3. Batas Ambang (WHO, 2003).....	9
Tabel 4. Data asli dan (*) data sintetik hasil SMOTE.....	15
Tabel 5. Hasil distribusi data sebelum dan sesudah SMOTE .....	17
Tabel 6. Kriteria AUC.....	22
Tabel 7. Data atribut.....	23
Tabel 8. Contoh Data .....	25
Tabel 9. Contoh data tingkat kemungkinan obesitas .....	27
Tabel 10. Klasifikasi tingkat kemungkinan obesitas .....	28
Tabel 11. Contoh data setelah <i>preprocessing</i> .....	29
Tabel 12. Pembagian dataset.....	30
Tabel 13. Contoh akurasi setiap fold.....	31
Tabel 14. Data <i>training</i> dan data <i>testing</i> .....	32
Tabel 15. Inisialisasi nilai bias .....	32
Tabel 16. Inisialisasi nilai bobot .....	32
Tabel 17. Matriks Hinit data training.....	33
Tabel 18. Matriks H data training .....	34
Tabel 19. Matriks hasil perkalian $H_t$ dengan H .....	34
Tabel 20. Invers matrix .....	34
Tabel 21. Matriks $H^+$ .....	35
Tabel 22. Nilai $\beta$ .....	35
Tabel 23. Matriks Hinit data testing.....	36
Tabel 24. Matriks H data testing .....	36
Tabel 25. Matriks $y$ .....	36
Tabel 26. Data prediksi dan Data target.....	36
Tabel 27. Jumlah kelas pada data training tanpa SMOTE.....	37
Tabel 28. Jumlah kelas pada data training dengan SMOTE .....	38

Tabel 29. Contoh data untuk perhitungan SMOTE .....	39
Tabel 30. menunjukkan hasil distribusi data sebelum dan sesudah SMOTE .....	40
Tabel 31. Hasil SMOTE data <i>training</i> .....	41
Tabel 32. Klasifikasi ELM tanpa SMOTE berdasarkan jumlah neuron .....	41
Tabel 33. Confusion matrix dengan 100 <i>hidden neuron</i> tanpa SMOTE.....	42
Tabel 34. Hasil presisi dan recall dengan 100 hidden neuron tanpa SMOTE .....	43
Tabel 35. Klasifikasi ELM dengan SMOTE.....	44
Tabel 36. Hasil confusion matrix dengan jumlah neuron 500 dengan SMOTE ..	45
Tabel 37. Hasil presisi dan recall berdasarkan 500 neuron dengan SMOTE.....	46
Tabel 38. Hasil klasifikasi berdasarkan jumlah <i>hidden neuron</i> .....	50

## DAFTAR GAMBAR

<b>Gambar</b>	<b>Halaman</b>
Gambar 1. Tahapan data mining .....	10
Gambar 2. Konsep kerja Stratified `10 fold cross validation.....	13
Gambar 3. Cara kerja SMOTE.....	15
Gambar 4. Struktur umum ELM .....	18
Gambar 5. Confusion Matrix .....	20
Gambar 6. Alur Penelitian.....	25
Gambar 7. Jumlah kelas pada data <i>training</i> tanpa SMOTE.....	38
Gambar 8. Jumlah kelas pada data training dengan SMOTE .....	39
Gambar 9. Kurva <i>multiclass</i> AUC 100 <i>hidden neuron</i> tanpa SMOTE.....	43
Gambar 10. kurva <i>multiclass</i> AUC 500 <i>hidden neuron</i> dengan SMOTE .....	46
Gambar 11. Perbandingan akurasi berdasarkan jumlah <i>hidden neuron</i> .....	47
Gambar 12. Perbandingan hasil klasifikasi menggunakan 100 <i>hidden neuron</i> .....	48
Gambar 13. Hasil akurasi berdasarkan <i>hidden neuron</i> .....	51

## DAFTAR LAMPIRAN

### **Lampiran**

Lampiran 1. *Source code dataset*

Lampiran 2. *Source code Preprocessing*

Lampiran 3. *Source Code k-fold*

Lampiran 4. Source code klasifikasi ELM

Lampiran 5. *Source code SMOTE*

Lampiran 6. Source code AUC

Lampiran 7. Source code menampilkan kurva multiclass AUC