



**STUDI KOMPARATIF VARIAN *SMOTE* PADA *DATASET* PENYAKIT
GINJAL KRONIK MENGGUNAKAN KLASIFIKASI *NAÏVE BAYES***

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Sarjana Strata-1 Ilmu Komputer**

**Oleh
KHAIRUL FURQAN
NIM. 1811016210017**

**PROGRAM STUDI S1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

JUNI 2025



**STUDI KOMPARATIF VARIAN *SMOTE* PADA *DATASET* PENYAKIT
GINJAL KRONIK MENGGUNAKAN KLASIFIKASI *NAÏVE BAYES***

Skripsi

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

Oleh

KHAIRUL FURQAN

NIM. 1811016210017

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

JUNI 2025

SKRIPSI

STUDI KOMPARATIF *SMOTE* VARIAN PADA *DATASET* PENYAKIT GINJAL KRONIK MENGGUNAKAN KLASIFIKASI *NAÏVE BAYES*

Oleh :

KHAIRUL FURQAN
1811016210017

Telah dipertahankan di depan Dosen Penguji pada tanggal 25 Juni 2025:

Susunan Dosen Penguji :

Pembimbing I



Rudy Herteno, S.Kom., M.Kom

NIP. 198809252022031003

Dosen Penguji I



Muliadi, S.Kom., M.Cs

NIP. 197804222010121002

Pembimbing II



Irwan Budiman, S.T., M.Kom

NIP. 197703252008121001

Dosen Penguji II



Setyo Wahyu Saputro, S.kom., M.Kom

NIP. 198808072023211027

Banjarbaru, 25 Juni 2025

Koordinator Program Studi Ilmu Komputer



Dwi Kartini, S.Kom., M.Kom

NIP. 198704212012122003

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, 25 Juni 2025



Khairul Furqan

NIM. 1811016210017

ABSTRAK

STUDI KOMPARATIF VARIAN *SMOTE* PADA *DATASET* PENYAKIT GINJAL KRONIK MENGGUNAKAN KLASIFIKASI *NAÏVE BAYES*

(Oleh: Khairul Furqan; Pembimbing: Rudy Herteno, S.Kom., M.Kom. dan Irwan Budiman, S.T., M.Kom. ; 2025; 73 halaman)

Penyakit Ginjal Kronik (PGK) merupakan salah satu penyakit kronis yang membutuhkan deteksi dini untuk penanganan yang tepat, karena seringkali Penyakit Ginjal Kronik (PGK) tidak terdeteksi hingga stadium lanjut karena gejala awal yang tidak spesifik. Keterlambatan diagnosis ini berakibat pada penanganan yang lebih sulit dan *prognosis* yang buruk. Oleh karena itu, pengetahuan mengenai kasus penyakit ginjal kronik sangat penting untuk memungkinkan penanganan awal yang dapat dicapai dengan menerapkan teknik klasifikasi. Namun, tantangan utama dalam klasifikasi penyakit ini adalah ketidakseimbangan data, di mana jumlah data pada kelas minoritas jauh lebih sedikit dibanding kelas mayoritas. Metode *Naïve Bayes* dipilih karena kesederhanaan dan efisiensinya dalam pengklasifikasian. Penelitian ini bertujuan untuk mengevaluasi kinerja metode klasifikasi *Naïve Bayes* pada *dataset* penyakit ginjal kronik dengan dan tanpa penggunaan metode *Oversampling Synthetic Minority Oversampling Technique (SMOTE)* beserta variannya. Proses penelitian meliputi tahap *preprocessing*, pembagian data menggunakan split 80:20 serta evaluasi dengan metrik akurasi, presisi, *recall*, dan *AUC*. Hasil pengujian menunjukkan bahwa penggunaan varian *SMOTE* mampu meningkatkan performa klasifikasi dibandingkan dengan tanpa *SMOTE*. Varian seperti *Supervised-SMOTE*, *Assembled-SMOTE*, dan *SMOTE Lee* yang mencapai akurasi hingga 0,9875 dan *AUC* di atas 0,98.

Kata Kunci : Penyakit Ginjal Kronik, *SMOTE*, Varian *SMOTE*, *Naïve Bayes* , Ketidakseimbangan Data

ABSTRACT

COMPARATIVE STUDY OF SMOTE VARIANTS ON CHRONIC KIDNEY DISEASE DATASET USING NAÏVE BAYES CLASSIFICATION

(By: Khairul Furqan; Mentor: Rudy Herteno, S.Kom., M.Kom. and Irwan Budiman, S.T., M.Kom. ; 2025; 73 page)

Chronic kidney disease (CKD) is a chronic condition that requires early detection for appropriate treatment, as it often goes undetected until advanced stages due to nonspecific early symptoms. This delayed diagnosis results in more complex treatment and poor prognosis. Therefore, knowledge and early identification of CKD cases are essential and can be supported through classification techniques. However, a major challenge in classifying CKD data is class imbalance, where the minority class contains significantly fewer instances than the majority class. The Naïve Bayes method was selected due to its simplicity and efficiency in classification tasks. This study aims to evaluate the performance of the Naïve Bayes classification algorithm on the CKD dataset, both with and without the application of the Synthetic Minority Oversampling Technique (SMOTE) and its variants. The research pROCeSS included data preprocessing, an 80:20 train-test split, and performance evaluation using accuracy, precision, recall, and AUC metrics. Experimental results indicate that the use of SMOTE variants enhances classification performance compared to models without SMOTE. Variants such as Supervised-SMOTE, Assembled-SMOTE, and SMOTE Lee achieved accuracy up to 0.9875 and AUC scores exceeding 0.98.

Keywords : Chronic kidney disease, SMOTE, SMOTE Variants, Naïve Bayes , Imbalanced Data

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 “Studi Komparatif Varian *SMOTE* Pada *Dataset* Ginjal Kronik Menggunakan Klasifikasi *Naïve Bayes*” 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 Munawir dan Ibu Rugayah yang selalu memberikan bantuan, semangat, doa, dan dukungan dalam proses menyelesaikan skripsi ini.
2. Bapak Rudy Herteno, S.Kom., M.Kom selaku dosen pembimbing utama yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
3. Bapak Irwan Budiman, S.T., M.Kom selaku dosen pembimbing pendamping yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian skripsi ini.
4. Ibu Dwi Kartini, S.Kom., M.Kom selaku Ketua Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan izin beliau skripsi ini dapat diselesaikan.
5. Bapak Muliadi, S.Kom., M.Cs dan Bapak Setyo Wahyu Saputro, S.Kom., M.Kom selaku dosen penguji yang telah memberikan masukan-masukkan dalam skripsi ini.
6. Bapak Friska Abadi, S.Kom., M.Kom selaku dosen pembimbing akademik dan seluruh Dosen dan staf Program Studi Ilmu Komputer FMIPA ULM atas ilmun dan bantuan yang diberikan selama ini yang sangat bermanfaat.

7. Teman-teman dan kerabat khususnya teman-teman Ilmu Komputer 2018, yang sudah ikut serta dalam memberikan dukungan, memberikan doa serta membantu mengerjakan penelitian ini.
8. 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, 25 Juni 2025



Khairul Furqan

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN.....	iii
ABSTRAK	iv
ABSTRACT	v
PRAKATA.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN.....	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan	3
1.4 Manfaat Penelitian	3
1.5 Batasan Masalah	3
BAB II TINJAUAN PUSTAKA.....	5
2.1 Kajian Terdahulu	5
2.2 Keaslian Penelitian.....	8
2.3 Landasan Teori.....	10
2.3.1 Penyakit Ginjal Kronik	10
2.3.2 <i>Data mining</i>	11
2.3.3 Imbalanced Data	12

2.3.4	Normalisasi <i>Min-Max</i>	13
2.3.5	<i>Synthetic Minority Oversampling Technique</i>	13
2.3.6	Varian <i>SMOTE</i>	15
2.3.7	<i>Naïve Bayes Classifier</i>	37
2.3.8	Evaluasi.....	39
BAB III METODE PENELITIAN		43
3.1	Alat Penelitian.....	43
3.2	Bahan Penelitian	43
3.3	Prosedur Penelitian	44
BAB IV HASIL DAN PEMBAHASAN		46
4.1	Hasil	46
4.1.1	Pengumpulan Data	46
4.1.2	<i>Preprocessing</i> Data.....	48
4.1.3	Klasifikasi	49
4.2	Pembahasan.....	63
BAB V PENUTUP.....		67
5.1	Kesimpulan	67
5.2	Saran	67
DAFTAR PUSTAKA		69

DAFTAR TABEL

Tabel	Halaman
Tabel 1. Keaslian Penelitian.....	8
Tabel 2. Rancangan Penelitian.....	9
Tabel 3. Kriteria AUC.....	42
Tabel 4. Contoh Data penyakit ginjal kronik	43
Tabel 5. Fitur penyakit ginjal kronik.....	46
Tabel 6. Data penyakit ginjal kronik.....	47
Tabel 7. Kelas penyakit ginjal kronik	47
Tabel 8. Contoh data setelah <i>preprocessing</i>	48
Tabel 9. Hasil klasifikasi <i>Naïve Bayes</i> tanpa <i>SMOTE</i>	49
Tabel 10. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>SMOTE</i>	50
Tabel 11. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>Lee</i>	51
Tabel 12. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>Supervised-SMOTE</i>	51
Tabel 13. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>Assembled-SMOTE</i>	52
Tabel 14. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>Borderline-SMOTE1</i>	53
Tabel 15. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>Borderline-SMOTE2</i>	53
Tabel 16. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>CURE -SMOTE</i>	54
Tabel 17. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>LN-SMOTE</i>	55
Tabel 18. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>Selected-SMOTE</i>	55
Tabel 19. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>SL-graph-SMOTE</i>	56
Tabel 20. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>SMOTE-OUT</i>	57
Tabel 21. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>SMOTEBW</i>	57
Tabel 22. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>SSO</i>	58
Tabel 23. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>SUNDO</i>	59
Tabel 24. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>ADOMS</i>	59
Tabel 25. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>DEAGO</i>	60
Tabel 26. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>Kmeans-SMOTE</i>	61
Tabel 27. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>CE-SMOTE</i>	61

Tabel 28. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>ISMOTE</i>	62
Tabel 29. Hasil klasifikasi <i>Naïve Bayes</i> dengan <i>LVQ-SMOTE</i>	63
Tabel 30. Rekapitulasi hasil Klasifikasi <i>Naïve Bayes</i> dengan <i>SMOTE</i> dan Tanpa <i>SMOTE</i>	64

DAFTAR GAMBAR

Gambar	Halaman
Gambar 1. Tahapan <i>data mining</i>	11
Gambar 2. Skema kinerja <i>SMOTE</i>	14
Gambar 3. <i>Confusion matrix</i>	40
Gambar 4. Alur Penelitian.....	44
Gambar 5. Kurva <i>ROC Naïve Bayes</i> tanpa <i>SMOTE</i>	49
Gambar 6. Kurva <i>ROC Naïve Bayes</i> dengan <i>SMOTE</i>	50
Gambar 7. Kurva <i>ROC Naïve Bayes</i> dengan <i>Lee</i>	51
Gambar 8. Kurva <i>ROC Naïve Bayes</i> dengan <i>Supervised-SMOTE</i>	52
Gambar 9. Kurva <i>ROC Naïve Bayes</i> dengan <i>Assembled-SMOTE</i>	52
Gambar 10 Kurva <i>ROC Naïve Bayes</i> dengan <i>Borderline-SMOTE1</i>	53
Gambar 11. Kurva <i>ROC Naïve Bayes</i> dengan <i>Borderline-SMOTE2</i>	54
Gambar 12. Kurva <i>ROC Naïve Bayes</i> dengan <i>CURE -SMOTE</i>	54
Gambar 13. Kurva <i>ROC Naïve Bayes</i> dengan <i>LN-SMOTE</i>	55
Gambar 14. Kurva <i>ROC Naïve Bayes</i> dengan <i>Selected-SMOTE</i>	56
Gambar 15. Kurva <i>ROC Naïve Bayes</i> dengan <i>SL-graph-SMOTE</i>	56
Gambar 16. Kurva <i>ROC Naïve Bayes</i> dengan <i>SMOTE-OUT</i>	57
Gambar 17. Kurva <i>ROC Naïve Bayes</i> dengan <i>SMOTEWB</i>	58
Gambar 18. Kurva <i>ROC Naïve Bayes</i> dengan <i>SSO</i>	58
Gambar 19. Kurva <i>ROC Naïve Bayes</i> dengan <i>SUNDO</i>	59
Gambar 20. Kurva <i>ROC Naïve Bayes</i> dengan <i>ADOMS</i>	60
Gambar 21. Kurva <i>ROC Naïve Bayes</i> dengan <i>DEAGO</i>	60
Gambar 22. Kurva <i>ROC Naïve Bayes</i> dengan <i>Kmeans-SMOTE</i>	61
Gambar 23. Kurva <i>ROC Naïve Bayes</i> dengan <i>CE-SMOTE</i>	62
Gambar 24. Kurva <i>ROC Naïve Bayes</i> dengan <i>ISMOTE</i>	62
Gambar 25. Kurva <i>ROC Naïve Bayes</i> dengan <i>LVQ-SMOTE</i>	63
Gambar 26. Grafik hasil Klasifikasi <i>Naïve Bayes</i> dengan <i>SMOTE</i> dan Tanpa <i>SMOTE</i>	65

DAFTAR LAMPIRAN

Lampiran

Lampiran 1. *Source code Import and read Dataset*

Lampiran 2. *Source code Preprocessing*

Lampiran 3. *Source Code Split Data*

Lampiran 4. *Source code Naïve Bayes tanpa SMOTE*

Lampiran 5. *Source code Naïve Bayes dengan SMOTE*

Lampiran 6. *Source code AUC*

Lampiran 7. *Source code menampilkan kurva ROC*