



**PENANGANAN KETIDAKSEIMBANGAN KELAS DENGAN TEKNIK
BAGGING PADA METODE *SUPPORT VECTOR MACHINE*, *LOGISTIC
REGRESSION*, DAN *K-NEAREST NEIGHBOR* UNTUK PREDIKSI
KEBERLANGSUNGAN HIDUP PASIEN GAGAL JANTUNG**

Skripsi

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

Oleh

ICHSAN PRADANA

1711016210007

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

MEI 2023



**PENANGANAN KETIDAKSEIMBANGAN KELAS DENGAN TEKNIK
BAGGING PADA METODE *SUPPORT VECTOR MACHINE*, *LOGISTIC
REGRESSION*, DAN *K-NEAREST NEIGHBOR* UNTUK PREDIKSI
KEBERLANGSUNGAN HIDUP PASIEN GAGAL JANTUNG**

Skripsi

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

**Oleh
ICHSAN PRADANA
1711016210007**

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

MEI 2023

SKRIPSI

**PENANGANAN KETIDAKSEIMBANGAN KELAS DENGAN TEKNIK
BAGGING PADA METODE SUPPORT VECTOR MACHINE, LOGISTIC
REGRESSION, DAN K-NEAREST NEIGHBOR UNTUK PREDIKSI
KEBERLANGSUNGAN HIDUP PASIEN GAGAL JANTUNG**

Oleh :

ICHSAN PRADANA

1711016210007

Telah dipertahankan di depan Dosen Penguji pada tanggal 30 Mei 2023

Susunan Dosen Penguji :

Pembimbing I



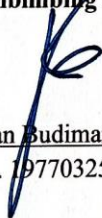
Triando Hamonangan Saragih, S.Kom., M.Kom
NIP. 199308242019031012

Dosen Penguji I



Muliadi, S.Kom., M.Sc
NIP. 197804222010121002

Pembimbing II



Irwan Budiman, S.T., M.Kom
NIP. 197703252008121001

Dosen Penguji II



Muhammad Itqan Mazdadi, M.Kom
NIP. 199006122019031013



Juni 2023

Program Studi Ilmu Komputer

Irwan 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 keserjanaan di suatu 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, Mei 2023

Yang Menyatakan,



Ichsan Pradana

NIM. 1711016210007

ABSTRAK

PENANGANAN KETIDAKSEIMBANGAN KELAS DENGAN TEKNIK BAGGING PADA METODE SUPPORT VECTOR MACHINE, LOGISTIC REGRESSION, DAN K-NEAREST NEIGHBOR UNTUK PREDIKSI KEBERLANGSUNGAN HIDUP PASIEN GAGAL JANTUNG (Oleh: Ichsan Pradana; Pembimbing: Triando Hamonangan Saragih, S.Kom., M.Kom. dan Irwan Budima, S.T., M.Kom.; 2023; 52 halaman)

Masalah ketidakseimbangan kelas pada dataset prediksi keberlangsungan hidup pasien gagal jantung merupakan tantangan yang perlu diatasi. Model prediksi cenderung memprioritaskan kelas mayoritas, mengakibatkan kesalahan prediksi yang berpotensi fatal. Dalam penelitian ini, teknik Bagging digunakan untuk mengatasi masalah ketidakseimbangan kelas. Bagging melakukan resampling dengan penggantian pada data training untuk meningkatkan representasi data kelas minoritas. Penelitian ini melibatkan tiga algoritma klasifikasi, yaitu *Support Vector Machine* (SVM), *Logistic Regression* (LR), dan *K-Nearest Neighbor* (KNN). Untuk meningkatkan kinerja algoritma-algoritma ini, dilakukan *hyperparameter tuning* pada *Bagging* dengan menggunakan kombinasi hyperparameter yang dihasilkan melalui algoritma *Grid Search*. Hasil penelitian menunjukkan bahwa penggunaan *Bagging* secara signifikan meningkatkan kinerja KNN yang memiliki varian tinggi. Pada SVM dan LR yang memiliki varian rendah, peningkatan kinerja oleh *Bagging* tidak terlalu signifikan. Hal ini dapat dijelaskan oleh keberagaman *data training* yang dihasilkan oleh *Bagging*, yang mengurangi varian pada algoritma dengan varian tinggi seperti KNN. Penanganan ketidakseimbangan kelas pada prediksi keberlangsungan hidup pasien gagal jantung melalui teknik *Bagging* ini memberikan pemahaman yang lebih baik mengenai pengaruhnya terhadap kinerja algoritma klasifikasi. Dengan mengurangi kesalahan prediksi, diharapkan penelitian ini dapat membantu dalam memberikan penanganan yang tepat bagi pasien gagal jantung dan meningkatkan hasil prediksi keberlangsungan hidup pasien secara keseluruhan.

Kata Kunci: Gagal Jantung, Ketidakseimbangan Kelas, *K-Nearest Neighbor*, *Support Vector Machine*, *Logistic Regression*, *Bagging*, *Grid Search*

ABSTRACT

HANDLING CLASS IMBALANCE WITH BAGGING TECHNIQUE USING SUPPORT VECTOR MACHINE, LOGISTIC REGRESSION, AND K-NEAREST NEIGHBOR METHODS FOR PREDICTING THE SURVIVAL OF HEART FAILURE PATIENTS (By: Ichsan Pradana; Supervisors: Triando Hamonangan Saragih, S.Kom., M.Kom. and Irwan Budiman, S.T., M.Kom.; 2023; 52 pages)

The class imbalance problem in predicting the survival of heart failure patients is a challenge that needs to be addressed. Predictive models tend to prioritize the majority class, leading to potentially fatal prediction errors. In this study, the Bagging technique is employed to tackle the class imbalance issue. Bagging performs resampling with replacement on the training data to improve the representation of the minority class. The study involves three classification algorithms: Support Vector Machine (SVM), Logistic Regression (LR), and K-Nearest Neighbor (KNN). To enhance the performance of these algorithms, hyperparameter tuning is conducted on Bagging using a combination of hyperparameters generated through the Grid Search algorithm. The results of the study demonstrate that the use of Bagging significantly improves the performance of KNN, which has high variance. For SVM and LR, which have low variance, the performance enhancement by Bagging is less significant. This can be attributed to the diversity of the training data generated by Bagging, which reduces variance in algorithms with high variance like KNN. Handling the class imbalance in predicting the survival of heart failure patients through the Bagging technique provides better understanding of its impact on the performance of classification algorithms. By reducing prediction errors, this research is expected to contribute to providing appropriate treatment for heart failure patients and improving overall predictions of patient survival.

Keywords: *Heart Failure, Class Imbalance, K-Nearest Neighbor, Support Vector Machine, Logistic Regression, Bagging, Grid Search.*

PRAKATA

Puji syukur penulis panjatkan kepada Allah SWT karena atas berkat, rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul **PENANGANAN KETIDAKSEIMBANGAN KELAS DENGAN TEKNIK BAGGING PADA METODE SUPPORT VECTOR MACHINE, LOGISTIC REGRESSION, DAN K-NEAREST NEIGHBOR UNTUK PREDIKSI KEBERLANGSUNGAN HIDUP PASIEN GAGAL JANTUNG** untuk memenuhi syarat dalam menyelesaikan pendidikan program S-1 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 kedua orang tua saya yang selalu memberikan do'a, semangat, bantuan dan dukungan dalam proses penyelesaian skripsi ini.
2. Bapak Triando Hamonangan Saragih, S.Kom., M.Kom. selaku dosen pembimbing utama yang senantiasa membimbing, membantu, meluangkan waktu dan memberikan energi positif demi kelancaran dalam proses penyelesaian skripsi ini.
3. Bapak Irwan Budiman, S.T., M.Kom. selaku dosen pembimbing pendamping yang turut serta memberikan arahan, membantu dan meluangkan waktu demi kelancaran dalam proses penyelesaian skripsi ini. Bapak Irwan Budiman, S.T., M.Kom. juga merupakan Koordinator Program Studi Ilmu Komputer FMIPA ULM, sehingga atas bantuan dan izin beliau skripsi ini dapat diselesaikan.
4. Seluruh dosen dan staf program studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama berkuliah di Ilmu Komputer yang sangat bermanfaat.

5. Teman – teman keluarga Ilmu Komputer angkatan 2017 yang memberikan bantuan dan dukungan yang senantiasa menjadi alat pacu penulis untuk selalu berusaha dan semangat untuk menyelesaikan skripsi.
6. Teman – teman yang telah banyak membantu dalam memberikan saran, motivasi dan yang selalu ada ketika aku berbagi keluh kesah selama proses penyelesaian skripsi.

Akhir kata penulis menyadari sepenuhnya bahwa penulisan skripsi ini masih jauh dari kesempurnaan, maka penulis mengharapkan bantuan berupa kritik dan saran yang membangun dari semua pihak sangat diharapkan demi penyempurnaan dan mutu penulisan skripsi ini. Semoga tulisan ini dapat bermanfaat bagi semua pihak untuk dapat menambah pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, Mei 2023



Ichsan Pradana
NIM. 1711016210007

DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PENGESAHAN	ii
PERNYATAAN.....	iii
ABSTRAK	iv
ABSTRACT	v
PRAKATA.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN.....	xiii
BAB I	
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	3
BAB II	
2.1 Kajian Terdahulu	4
2.2 Landasan Teori	6
2.2.1 Gagal Jantung	6
2.2.2 Ketidakseimbangan Kelas	6
2.2.3 Dataset <i>Heart Failure Clinical Records</i>	6
2.2.4 <i>Bagging</i>	8
2.2.5 <i>Hyperparameter Tuning</i>	9
2.2.6 <i>Grid Search</i>	9
2.2.7 <i>Support Vector Machine</i>	9
2.2.8 <i>Logistic Regression</i>	10
2.2.9 <i>K-Nearest Neighbor</i>	11
2.2.10 <i>Cross Validation</i>	12

2.2.11 <i>Confusion Matrix</i>	12
2.2.12 <i>Area Under the ROC (Receiver Operating Characteristic) Curve</i>	13
BAB III	
3.1 Alat Penelitian	15
3.2 Bahan Penelitian	15
3.3 Variabel Penelitian.....	15
3.4 Prosedur Penelitian	15
BAB IV	
4.1 Hasil.....	19
4.1.1 Pengumpulan Data	19
4.1.2 Analisis Ketidakseimbangan Kelas.....	21
4.1.3 <i>Normalisasi Data</i>	21
4.1.4 <i>Split Data</i>	23
4.1.5 Klasifikasi tanpa <i>Bagging</i>	24
4.1.6 <i>Hyperparameter Tuning</i>	35
4.1.7 Klasifikasi dengan <i>Bagging</i>	45
4.2 Pembahasan	45
BAB V	
5.1 Kesimpulan.....	49
5.2 Saran	49
DAFTAR PUSTAKA	50
LAMPIRAN.....	52

DAFTAR TABEL

Tabel	Halaman
Tabel 1. Keaslian Penelitian.....	5
Tabel 2. Tabel Rancangan Penelitian.....	5
Tabel 3. Rincian dataset heart Failure Clinical Records.....	7
Tabel 4. Sampel dataset <i>Heart Failure Clinical Records</i>	7
Tabel 5. <i>Hyperparameter</i> pada <i>bagging</i>	8
Tabel 6. Kualitas klasifikasi berdasarkan AUC.....	14
Tabel 7. Hyperparameter Bagging.....	18
Tabel 8. Fitur dataset <i>Heart Failure Clinical Records</i>	19
Tabel 9. Contoh dataset <i>Heart Failure Clinical Records</i>	20
Tabel 10 Perhitungan Manual Normalisasi Data.....	21
Tabel 11. Dataset.....	22
Tabel 12. Nilai Kinerja KNN tanpa <i>bagging</i>	24
Tabel 13. Contoh Data Testing Perhitungan Manual KNN.....	24
Tabel 14. Data Training Perhitungan Manual KNN.....	24
Tabel 15. Perhitungan Jarak KNN.....	25
Tabel 16. Confusion Matriks perhitungan KNN.....	25
Tabel 17. Nilai Kinerja SVM tanpa <i>bagging</i>	26
Tabel 18. Contoh <i>Data Testing</i> Perhitungan Manual SVM.....	26
Tabel 19. Data Training Perhitungan Manual SVM.....	27
Tabel 20. Koefisien setiap kolom SVM (X) dan intercept (C).....	27
Tabel 21. Contoh Perhitungan data testing SVM.....	28
Tabel 22. Prediksi keseluruhan data testing SVM.....	29
Tabel 23. <i>Confusion Matrix</i> perhitungan SVM.....	30
Tabel 24. Nilai Kinerja LR tanpa <i>bagging</i>	31
Tabel 25. Contoh <i>Data Testing</i> Perhitungan Manual <i>Logistic Regression</i>	31
Tabel 26. Data Training Perhitungan Manual <i>Logistic Regression</i>	31
Tabel 27. Koefisien setiap kolom LR (X) dan <i>Intercept</i> (C).....	32

Tabel 28. Prediksi pada seluruh data testing LR.....	32
Tabel 29. <i>Confusion Matrix</i> perhitungan LR.....	34
Tabel 30. <i>Hyperparameter tuning</i> bagging-KNN : AUC.....	35
Tabel 31 Hyperparameter Tuning Bagging-KNN : <i>Accuracy</i>	36
Tabel 32 Hyperparameter Tuning Bagging-KNN : <i>Precision</i>	36
Tabel 33. Hyperparameter Tuning Bagging-KNN : <i>Recall</i>	37
Tabel 34. Hyperparameter Tuning Bagging-KNN : <i>F1 Score</i>	38
Tabel 35. <i>Hyperparameter tuning</i> bagging-SVM : AUC.....	38
Tabel 36. Hyperparameter Tuning Bagging-SVM : <i>Accuracy</i>	39
Tabel 37. Hyperparameter Tuning Bagging-SVM : <i>Precision</i>	39
Tabel 38. Hyperparameter Tuning Bagging-SVM : <i>Recall</i>	40
Tabel 39. Hyperparameter Tuning Bagging-SVM : <i>F1 Score</i>	41
Tabel 40. <i>Hyperparameter tuning</i> bagging-LR : AUC.....	41
Tabel 41. <i>Hyperparameter tuning</i> bagging-LR : <i>Accuracy</i>	42
Tabel 42. <i>Hyperparameter tuning</i> bagging-LR : <i>Precision</i>	42
Tabel 43. <i>Hyperparameter tuning</i> bagging-LR : <i>Recall</i>	43
Tabel 44. <i>Hyperparameter tuning</i> bagging-LR : <i>F1 Score</i>	44
Tabel 45. Hyperparameter Terbaik.....	44
Tabel 46. Pengujian <i>Hyperparameter Bagging</i>	45
Tabel 47. Perbandingan Klasifikasi dengan <i>Bagging</i> dan Tanpa <i>Bagging</i>	45

DAFTAR GAMBAR

Gambar	Halaman
Gambar 1 <i>Bagging</i>	8
Gambar 2 <i>Grid Search</i>	9
Gambar 3 <i>Support Vector Machine</i>	10
Gambar 4 <i>Logistic Regresion</i>	11
Gambar 5 <i>Stratified 10-Fold Cross Validation</i>	12
Gambar 6 <i>Counfusion Matrix</i>	13
Gambar 7 Kurva ROC.....	14
Gambar 8 Alur Penelitian.....	16
Gambar 9 Perbandingan persentase kelas <i>deceased</i> dan <i>nondeceased</i>	21
Gambar 10. Perbandingan jumlah data training dan testing.....	23

DAFTAR LAMPIRAN

Lampiran

Lampiran 1. Dataset Heart Failure Clinical Records

Lampiran 2. *Estimator Samples & Features Bagging*

Lampiran 3. Contoh Prediksi dengan *Bagging*

Lampiran 4. *Data Training Scaled*

Lampiran 5. *Data Testing Scaled*

Lampiran 6. Daftar Hidup Penulis