

**PERBAIKAN *PERFORMANCE* ALGORITMA C4.5 ADABOOST
BERBASIS OPTIMASI PROSES SELEKTIF FITUR**

TESIS



**OLEH
GEDE AJUS SETIAWAN
1829101001**

**UNIVERSITAS PENDIDIKAN GANESHA
PROGRAM PASCASARJANA
PROGRAM STUDI ILMU KOMPUTER**

2023

PERBAIKAN *PERFORMANCE* ALGORITMA C4.5 ADABOOST BERBASIS
OPTIMASI PROSES SELEKTIF FITUR

TESIS



OLEH
GEDE AJUS SETIAWAN
1829101001

UNIVERSITAS PENDIDIKAN GANESHA
PROGRAM PASCASARJANA
PROGRAM STUDI ILMU KOMPUTER

2023

Lembar Persetujuan Pembimbing

Tesis oleh Gede Ajus Setiawan dengan judul Perbaikan Performance Algoritma C4.5 Adaboost Berbasis Optimasi Proses Selektif Fitur telah diperiksa dan disetujui untuk Mengikuti Ujian Tesis.

Singaraja, 17 Juli 2023



Pembimbing I

Dr. I Gede Aris Gunadi, S.Si., M.Kom.
NIP 197703182008121004



Pembimbing II

Dr. Gede Indrawan, S.T., M.T.
NIP 197601022003121001

Lembar Persetujuan Tim Penguji

Tesis oleh Gede Ajus Setiawan ini telah dipertahankan di depan tim penguji dan dinyatakan diterima sebagai salah satu persyaratan untuk memperoleh gelar Magister Komputer di Program Studi Ilmu Komputer, Program Pascasarjana, Disetujui pada tanggal : 17 Juli 2023

Oleh Tim Penguji



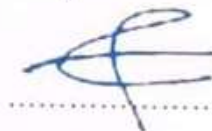
Ketua Dr. I Made Gede Sunarya, S.Kom., M.Cs.
NIP 198307252008011008



Anggota Dr. I Gede Aris Gunadi, S.Si., M.Kom.
NIP 197703182008121004



Anggota Dr. Gede Indrawan, S.T., M.T.
NIP 197601022003121001



Anggota Prof. Dr. Sariyasa, M.Sc., Ph.D.
NIP 196512291990032002



Anggota Dr. Gede Suweken, M.Sc.
NIP 196710131994031001

Mengetahui Direktur,

Program Pascasarjana Undiksha,

Prof. Dr. Nyoman Jampel, M.Pd.
NIP 19591010198603100

LEMBAR PERNYATAAN

Saya menyatakan dengan sesungguhnya, bahwa tesis yang saya susun sebagai syarat untuk memperoleh gelar Magister Komputer dari Program Pascasarjana Universitas Pendidikan Ganesha seluruhnya merupakan hasil karya saya sendiri. Bagian-bagian tertentu dalam penulisan tesis yang saya kutip dari hasil karya orang lain telah dituliskan sumbernya secara jelas dan sesuai dengan norma, kaidah, serta etika akademis.

Apabila dikemudian hari ditemukan seluruh atau sebagian tesis ini bukan hasil karya saya sendiri atau adanya plagiat dalam bagian-bagian tertentu, saya bersedia menerima sanksi pencabutan gelar akademik yang saya sandang dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku di wilayah Negara Kesatuan Republik Indonesia.

Singaraja, 17 Juli 2023

Yang memberi pernyataan



Gede Ajus Setiawan

NIM: 1829101001

PRAKATA

Puji dan syukur penulis panjatkan kehadapan Tuhan Yang Maha Esa karena atas anugrah-Nya, sehingga tesis yang berjudul Perbaikan *Performance* Algoritma C4.5 Adaboost Berbasis Optimasi Proses Selektif Fitur bisa terselesaikan dengan tepat waktu.

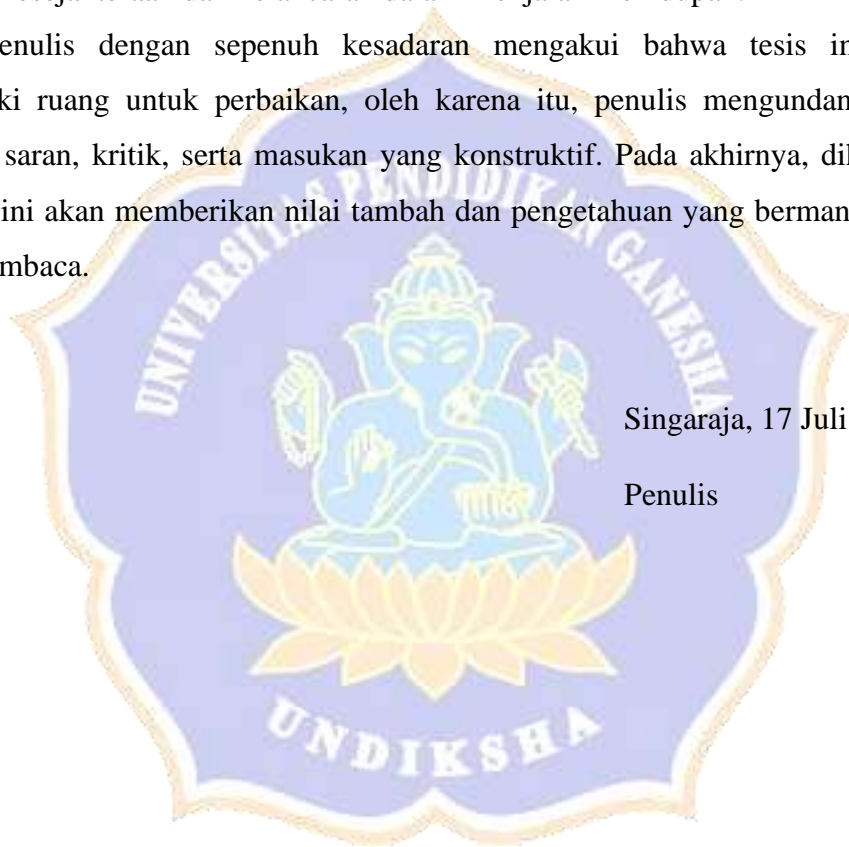
Tesis ini disusun dengan tujuan memenuhi salah satu syarat untuk meraih gelar Magister Ilmu Komputer di Program Pascasarjana Universitas Pendidikan Ganesha. Penyelesaian tesis ini didukung oleh banyak individu dan organisasi yang telah memberikan kontribusi yang berharga. Oleh karena itu, kami ingin mengungkapkan rasa terima kasih dan penghargaan yang tulus kepada semua yang telah berperan.

1. Dr. I Gede Aris Gunadi, S.Si., M.Kom, sebagai pembimbing I yang telah dengan sabar membimbing, mengarahkan, dan memberikan motivasi yang demikian bermakna, sehingga penulis mampu melewati berbagai hambatan dalam perjalanan studi dan penyelesaian tesis ini;
2. Dr. Gede Indrawan, S.T., M.T, sebagai pembimbing II, yang juga dengan sabar membimbing, mengarahkan, dan memberikan motivasi selama penyusunan tesis sehingga dapat terwujud dengan baik sesuai harapan;
3. Prof. Dr. Sariyasa, M.Sc., Ph.D. dan Dr. Gede Suweken, M.Sc. sebagai penguji yang telah banyak memberikan saran-saran yang bermanfaat untuk menyempurnakan tesis ini;
4. Kepala UPT-TIK Undiksha yang telah memberikan izin dan bantuan administratif selama penyusunan tesis ini;
5. Ketua Program Studi Ilmu Komputer dan staf dosen pengajar di program studi Ilmu Komputer yang sudah banyak memberikan motivasi dan memberikan bantuan penulis selama menyusun tesis ini;
6. Direktur Program Pascasarjana Undiksha dan staf, yang telah banyak membantu selama penulis menyelesaikan tesis ini;

7. Rektor Universitas Pendidikan Ganesha, yang telah memberikan bantuan secara moral dan memfasilitasi berbagai kepentingan penulis dalam menyelesaikan tesis ini;
8. Rekan-rekan seangkatan di Program Studi Ilmu Komputer yang sudah banyak memberikan dukungan untuk menciptakan kemandirian peneliti selama menempuh pendidikan serta menuntaskan tesis ini;

Semoga segala dukungan yang sudah diberikan dalam penyelesaian penelitian ini akan mendapatkan balasan setimpal dari Tuhan Yang Maha Esa berupa kesejahteraan dan kelancaran dalam menjalani kehidupan.

Penulis dengan sepenuh kesadaran mengakui bahwa tesis ini masih memiliki ruang untuk perbaikan, oleh karena itu, penulis mengundang segala bentuk saran, kritik, serta masukan yang konstruktif. Pada akhirnya, diharapkan tulisan ini akan memberikan nilai tambah dan pengetahuan yang bermanfaat bagi para pembaca.



Singaraja, 17 Juli 2023

Penulis

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL.....	ii
Lembar Persetujuan Pembimbing	iii
Lembar Persetujuan Tim Penguji.....	iv
LEMBAR PERNYATAAN	v
PRAKATA.....	vi
ABSTRAK	viii
ABSTRACT.....	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR	xii
DAFTAR TABEL.....	xv
DAFTAR LAMPIRAN.....	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan Masalah.....	4
1.4 Batasan Masalah	4
1.5 Tujuan Penelitian	5
1.6 Manfaat Penelitian	5
BAB II KAJIAN PUSTAKA.....	7
2.1 Kajian Teori.....	7
2.1.1 Data Mining.....	7
2.1.2 Clustering.....	8
1. Hierarchical clustering.....	12
2. Partitional Clustering	12
2.1.3 Selektif Fitur	12
2.1.4 Algoritma C4.5.....	18
2.1.5 Algoritma AdaBoost.....	20
2.1.6 Algoritma C4.5 dan AdaBoost.....	22

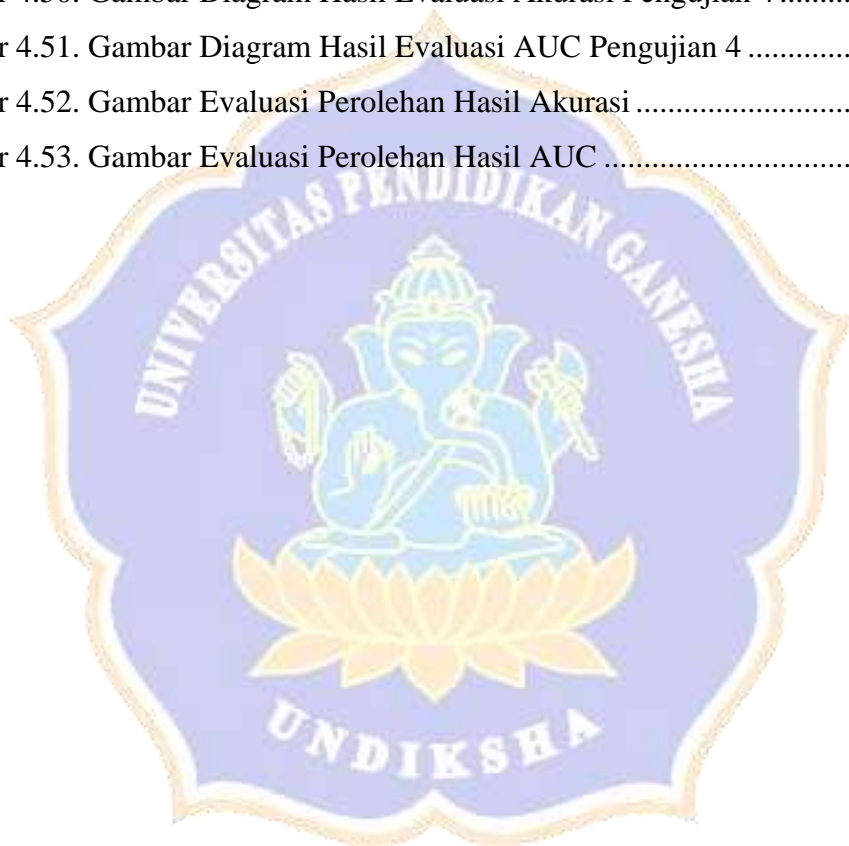
2.1.7	Confusion Matrix.....	23
2.2	Kajian Hasil Penelitian yang Relevan.....	24
2.3	Kerangka Konsep.....	29
BAB III METODOLOGI PENELITIAN.....		31
3.1	Prosedur Penelitian.....	31
3.1.1	Data Preprocessing.....	32
3.1.2	Model Algoritma.....	48
3.1.3	Evaluasi.....	50
BAB IV HASIL DAN PEMBAHASAN.....		53
4.1	Hasil.....	53
4.1.1	Pengujian 1 (Pengujian Algoritma C4.5).....	53
4.1.2	Pengujian 2 (Pengujian Algoritma C4.5 Berbasis AdaBoost).....	61
4.1.3	Pengujian 3 (Pengujian Algoritma C4.5 dengan Selektif Fitur).....	68
4.1.4	Pengujian 4 (Pengujian Algoritma C4.5 berbasis AdaBoost dengan Selektif Fitur).....	88
4.2	Pembahasan.....	91
BAB V PENUTUP.....		97
5.1	Simpulan.....	97
5.2	Saran.....	99
DAFTAR PUSTAKA.....		100
LAMPIRAN-LAMPIRAN.....		102
RIWAYAT HIDUP.....		135

DAFTAR GAMBAR

Gambar 2.1 Gambar Kerangka Berfikir.....	29
Gambar 3.1. Gambar Metode yang Diusulkan	31
Gambar 4.1. Gambar Model Pengujian 1 (Pengujian Algoritma C4.5).....	53
Gambar 4.2. Gambar Sub Proses <i>Cross Validation</i>	54
Gambar 4.3. Gambar Akurasi <i>Airline Passenger Satisfaction</i> Pengujian 1.....	55
Gambar 4.4. Gambar <i>ROC Curve Airline Passenger Satisfaction</i> Pengujian 1 ...	55
Gambar 4.5. Gambar Hasil Klasifikasi <i>Airline Passenger Satisfaction</i> Pengujian 1	56
Gambar 4.6. Gambar Akurasi <i>IRIS</i> pengujian 1.	56
Gambar 4.7. Gambar <i>ROC Curve IRIS</i> Pengujian 1	57
Gambar 4.8. Gambar Hasil Klasifikasi <i>IRIS</i> Pengujian 1	57
Gambar 4.9. Gambar Akurasi <i>TIC TAC TOE</i> pengujian 1.	58
Gambar 4.10. Gambar <i>ROC Curve TIC TAC TOE</i> Pengujian 1	58
Gambar 4.11. Gambar Hasil Klasifikasi <i>TIC TAC TOE</i> Pengujian 1	59
Gambar 4.12. Gambar Akurasi <i>Water_Potability</i> pengujian 1.....	59
Gambar 4.13. Gambar <i>ROC Curve Water_Potability</i> Pengujian 1.....	60
Gambar 4.14. Gambar Hasil Klasifikasi <i>Water_Potability</i> Pengujian 1.....	60
Gambar 4.15. Gambar Diagram Hasil Evaluasi Pengujian 1	61
Gambar 4.16. Gambar Model Pengujian 2 (Pengujian Algoritma C4.5 Berbasis AdaBoost)	62
Gambar 4.17. Gambar Sub Proses <i>Cross Validation</i> Pengujian 2.....	62
Gambar 4.18. Gambar Sub Proses AdaBoost Pengujian 2	62
Gambar 4.19. Gambar Akurasi <i>Airline Passenger Satisfaction</i> Pengujian 2.....	63
Gambar 4.20. Gambar <i>ROC Curve Airline Passenger Satisfaction</i> Pengujian 2..	64
Gambar 4.21. Gambar Akurasi <i>IRIS</i> pengujian 2.	64
Gambar 4.22. Gambar <i>ROC Curve IRIS</i> Pengujian 2	65
Gambar 4.23. Gambar Akurasi <i>TIC TAC TOE</i> pengujian 2.	65
Gambar 4.24. Gambar <i>ROC Curve TIC TAC TOE</i> Pengujian 2	66
Gambar 4.25. Gambar Akurasi <i>Water_Potability</i> pengujian 2.....	66
Gambar 4.26. Gambar <i>ROC Curve Water_Potability</i> Pengujian 2.....	67

Gambar 4.27. Gambar Diagram Hasil Evaluasi Pengujian 2.....	68
Gambar 4.28. Gambar Model Selektif Fitur PCA pada Dataset <i>Airline Passenger Satisfaction</i>	69
Gambar 4.29. Gambar Model Selektif Fitur <i>Info Gain</i> pada Dataset <i>Airline Passenger Satisfaction</i>	70
Gambar 4.30. Gambar Model Selektif Fitur <i>Chi-Square</i> pada Dataset <i>Airline Passenger Satisfaction</i>	71
Gambar 4.31. Gambar Model Selektif Fitur <i>Forward Selection</i> pada Dataset <i>Airline Passenger Satisfaction</i>	71
Gambar 4.32. Gambar Model Selektif Fitur <i>Backward Elemination</i> pada Dataset <i>Airline Passenger Satisfaction</i>	72
Gambar 4.33. Gambar Model Selektif Fitur PCA pada Dataset <i>IRIS</i>	73
Gambar 4.34. Gambar Model Selektif Fitur <i>Info Gain</i> pada Dataset <i>IRIS</i>	74
Gambar 4.35. Gambar Model Selektif Fitur <i>Chi Square</i> pada Dataset <i>IRIS</i>	75
Gambar 4.36. Gambar Model Selektif Fitur <i>Forward Selection</i> pada Dataset <i>IRIS</i>	75
Gambar 4.37. Gambar Model Selektif Fitur <i>Backward Elemination</i> pada Dataset <i>IRIS</i>	76
Gambar 4.38. Gambar Model Selektif Fitur PCA pada Dataset <i>TIC TAC TU</i>	77
Gambar 4.39. Gambar Model Selektif Fitur <i>Info Gain</i> pada Dataset <i>TIC TAC TU</i>	78
Gambar 4.40. Gambar Model Selektif Fitur <i>Chi Square</i> pada Dataset <i>TIC TAC TU</i>	79
Gambar 4.41. Gambar Model Selektif Fitur <i>Forward Selection</i> pada Dataset <i>TIC TAC TU</i>	80
Gambar 4.42. Gambar Model Selektif Fitur <i>Backward Elemination</i> pada Dataset <i>TIC TAC TU</i>	80
Gambar 4.43. Gambar Model Selektif Fitur PCA pada Dataset <i>Water_Potability</i>	81
Gambar 4.44. Gambar Model Selektif Fitur <i>Info Gain</i> pada Dataset <i>Water_Potability</i>	82

Gambar 4.45. Gambar Model Selektif Fitur <i>Chi Square</i> pada Dataset <i>Water_Potability</i>	83
Gambar 4.46. Gambar Model Selektif Fitur <i>Forward Selection</i> pada Dataset <i>Water_Potability</i>	84
Gambar 4.47. Gambar Model Selektif Fitur <i>Backward Elemination</i> pada Dataset <i>Water_Potability</i>	85
Gambar 4.48. Gambar Diagram Hasil Evaluasi Akurasi Pengujian 3	86
Gambar 4.49. Gambar Diagram Hasil Evaluasi AUC Pengujian 3	88
Gambar 4.50. Gambar Diagram Hasil Evaluasi Akurasi Pengujian 4	89
Gambar 4.51. Gambar Diagram Hasil Evaluasi AUC Pengujian 4	91
Gambar 4.52. Gambar Evaluasi Perolehan Hasil Akurasi	93
Gambar 4.53. Gambar Evaluasi Perolehan Hasil AUC	96



DAFTAR TABEL

Tabel 2.1. Tabel <i>Confusion Matrix</i>	23
Tabel 2.2. Tabel Penelitian Terkait	25
Tabel 2.3. Tabel Perlakuan Parameter Dataset	30
Tabel 3.1. Tabel Informasi Dataset yang Digunakan dalam Penelitian.	32
Tabel 3.2. Tabel Informasi Dataset <i>Airline Passenger Satisfaction</i>	32
Tabel 3.3. Tabel Contoh Dataset <i>Airline Passenger Satisfaction</i> (Kaggle Inc., 2023).....	36
Tabel 3.4. Tabel Informasi Dataset <i>IRIS</i>	38
Tabel 3.5. Tabel Contoh Dataset <i>Iris</i> (Kaggle Inc., 2023).....	39
Tabel 3.6. Tabel Informasi Dataset Tic-Tac-Toe.....	39
Tabel 3.7. Tabel Contoh Dataset <i>Tic-Tac-Toe</i> Sebelum <i>Preprocessing</i> (Kaggle Inc., 2023).....	41
Tabel 3.8. Tabel Contoh Dataset Tic-Tac-Toe sesudah <i>Preprocessing</i>	42
Tabel 3.9. Tabel Informasi Dataset <i>Water_Potability</i>	43
Tabel 3.10. Tabel Contoh Dataset <i>Water_Potability</i> Sebelum <i>Preprocessing</i> (Kaggle Inc., 2023).....	47
Tabel 3.11. tabel Contoh Dataset <i>Water_Potability</i> Sesudah <i>Preprocessing</i>	48
Tabel 3.12 Tabel Data <i>Dummy</i>	48
Tabel 3.13. Tabel Data Uji.....	50
Tabel 3.14. Tabel Hasil pengujian kedua algoritma pada data uji.....	51
Tabel 4.1. Tabel Hasil Evaluasi Pengujian 1	61
Tabel 4.2. Tabel Hasil Evaluasi Pengujian 2	67
Tabel 4.3. Tabel Hasil Evaluasi akurasi Pengujian 3 Algoritma C4.5 dengan Selektif Fitur	86
Tabel 4.4. Tabel Hasil Evaluasi AUC Pengujian 3 Algoritma C4.5 dengan Selektif Fitur	87
Tabel 4.5. Tabel Hasil Evaluasi Akurasi Pengujian 4 Algoritma C4.5 berbasis AdaBoost dengan Selektif Fitur	89
Tabel 4.6. Tabel Hasil Evaluasi AUC Pengujian 4 Algoritma C4.5 berbasis AdaBoost dengan Selektif Fitur.	90

Tabel 4.7. Tabel Evaluasi Perolehan Hasil Akurasi..... 92
Tabel 4.8. Tabel Evaluasi Perolehan Hasil AUC..... 94



DAFTAR LAMPIRAN

Lampiran 1. Informasi Dataset yang Digunakan	103
Lampiran 2. Hasil Perhitungan Manual	108
Lampiran 3. Hasil Perhitungan Dengan Sistem	130

