

**RANDOM FOREST DAN MULTIVARIATE ADAPTIVE
REGRESSION SPLINE UNTUK KLASIFIKASI
PENERIMAAN BEASISWA GENIUS S1 BAZNAS
KABUPATEN BOJONEGORO**

SKRIPSI

Disusun sebagai salah satu syarat
untuk memperoleh gelar Sarjana Statistika
Program Studi Statistika dan Sains Data



UNUGIRI

oleh

Iva Nur Achyani

2520190048

UNUGIRI

**PROGRAM STUDI STATISTIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS NAHDLATUL ULAMA SUNAN GIRI**

2023

PERYATAAN KEASLIAN TULISAN

Saya menyatakan bahwa skripsi ini bebas plagiat, dan apabila di kemudian hari terbukti terdapat plagiat dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan peraturan perundang-undangan.



HALAMAN PERSETUJUAN

Nama : Iva Nur Achyani

NIM : 2520190048

Judul : *Random Forest Dan Multivariate Adaptive Regression Spline Untuk Klasifikasi Penerimaan Beasiswa Genius S1 Baznas kabupaten Bojonegoro.*

Telah disetujui dan dinyatakan memenuhi syarat untuk diajukan dalam ujian skripsi.

Bojonegoro, 28 Juni 2023

Pembimbing I



Alif Yuanita Kartini, M.Si.

NIDN:0721048606

Pembimbing II

UNUGIRI

A large, stylized signature in black ink is centered over the text 'UNUGIRI'.

Fetrika Anggraini, M.Pd.

NIDN:0718038803

HALAMAN PENGESAHAN

Nama : Iva Nur Achyani

Nim : 2520190048

Judul : *Random Forest Dan Multivariate Adaptive Regression Spline Untuk Klasifikasi Penerimaan Beasiswa Genius S1 Baznas kabupaten Bojonegoro.*

Telah dipertahankan di hadapan penguji pada tanggal, 31 Juli 2023

Dewan Penguji

Penguji I



Nita Cahyani, M.Stat

NIDN: 0704038906

Penguji II



Dr. Hj. Ifa Khoiria-Ningrum, S.E., M.M.

NIDN: 0709097805

Tim Pembimbing

Pembimbing I



Alif Yuanita Kartini, M.Si.

NIDN:0721048606

Pembimbing II



Fetrika Angraini, M.Pd.

NIDN:0718038803

Mengetahui,

Dekan Fakultas Sains dan Teknologi



Sunu Wahyudhi, M.Pd.

NIDN:0709058902

Mengetahui,

Ketua Program Studi Statistika



Nita Cahyani, M.Stat.

NIDN:0704038906

HALAMAN MOTTO DAN PERSEMBAHAN

MOTTO

“So verily with the hardship there is relief, verily with the hardship there is relief. – (Q.S Al-Insyirah: 5-6)”

“Untuk masa sulitmu, biarlah Allah yang menguatkanmu. Tugas hanya berusaha agar jarak kamu dengan Allah tidak terlalu jauh”

“Tinggalkan apa yang ingin kamu tinggalkan, lakukanlah sesukamu dengan kerja kerasmu selagi itu tidak merugikan dirimu”

“Orang lain ga akan paham struggle dan masa sulitnya kita, yang mereka tahu hanya bagian success storiesnya aja. Jadi berjuanglah untuk diri sendiri meskipun ga ada yang tepuk tangan. Kelak diri kita akan bangga dengan apa yang kita perjuangkan hari ini. Tetap semangat ya memperjuangkan hidup yang penuh lika liku ini!”

PERSEMBAHAN

Pertama untuk diri saya sendiri, terimakasih sudah mau berjuang dan bertahan sampai saat ini sampai mampu berda pada titik ini.

Kedua untuk orang tua tercinta, ayahanda yang telah berada di surga dan ibu yang telah memberikan kasih sayang, dukungan dan yang merawat saya hingga titik ini.

Dan ucapan terimakasih kepada sahabat, kakak, adek, dosen, teman dan semua pihak yang telah bertanya “sampai mana skripsimu?”, “kapan wisuda” dan lain sebagainya. Sekali lagi terimakasih berkat kalian saya bisa sampai pada titik ini. Love you kalian semua ♡

KATA PENGANTAR

Puji dan syukur kehadirat Allah SWT karena atas limpahan rahmat, ridho, dan karunia-Nya sehingga dapat menyelesaikan skripsi yang berjudul “*Random Forest Dan Multivariate Adaptive Regression Spline Untuk Klasifikasi Penerimaan Beasiswa Genius S1 Baznas kabupaten Bojonegoro*” dengan lancar dan tepat waktu. Skripsi ini disusun untuk memenuhi persyaratan akademik untuk memperoleh gelar Sarjana pada Program Studi S1 Statistika Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.

Kelancaran dalam pengerjaan dan penyusunan Skripsi ini tidak terlepas dari bantuan berbagai pihak serta doa dan dukungan dari orang tua, keluarga, dan teman-teman. Dengan ketulusan hati, penulis mengucapkan terimakasih kepada:

1. Bapak K. M. Jauharul Ma’arif, M.Pd.I. selaku Rektor Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
2. Ibu Nita Cahyani, M.Stat. selaku Ketua Program Studi Statistika Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
3. Ibu Alif Yuanita Kartini, M.Si. selaku Dosen pembimbing satu yang sudah memberikan arahan dan dukungan serta membimbing hingga penulis dapat menyelesaikan skripsi dengan baik dan lancar.
4. Ibu Fetrika Anggraini, M.Pd. selaku Dosen pembimbing kedua yang sudah memberikan arahan dan selalu meyempatkan waktu untuk bimbingan.
5. Seluruh Dosen Program Studi Statistika Universitas Nahdlatul Ulama Sunan Giri Bojonegoro yang sudah memberikan serta mengajarkan banyak ilmu selama empat tahun.
6. Orang Tua serta Keluarga yang selalu mendo’akan serta memberikan semangat dengan tanpa lelah, sehingga penulis dapat menyelesaikan skripsi ini dengan mudah, lancar, dan lulus tepat waktu, tak lupa juga terimakasih sudah membiayai kuliah penulis dengan susah payah.
7. Teman-teman Program Studi Statistika dan Sains Data Universitas Nahdlatul Ulama Sunan Giri Bojonegoro angkatan 2019, serta sahabat dan adek tingkat yang selalu memberikan suport dan semangat tanpa lelah.
8. Teman-teman Baznas kabupaten Bojonegoro yang selalu memberi suport

positif sehingga peneliti sampai ditahap ini.

Penulis menyadari bahwa Skripsi ini masih jauh dari kata sempurna dan masih banyak kekurangan lainnya, maka dari itu penulis mengharapkan saran dan kritik yang membangun dari semua pihak demi kesempurnaan penulisan skripsi ini. Akhir kata penulis berharap semoga skripsi ini dapat bermanfaat bagi pembaca.

Bojonegoro, 21 Februari 2023

Penulis



UNUGIRI

ABSTRACT

Achyani, N, I. (2023). Random Forest and Multivariate Adaptive Regression Spline (MARS) for Classification of Genius S1 Baznas Scholarships, Bojonegoro district. Thesis. Statistics Study Program. Nahdlatul Ulama University Sunan Giri Bojonegoro. Main Supervisor Alif Yuanita Kartini, M.Sc. and second supervisor Fetrika Anggraini, M.Pd. Education problem in Indonesia is currently caused by the low quality of education. Bojonegoro Regency is still said to have low education quality, this can be seen in the School Participation Rate (APS) and Pure Participation Rate (APM) in Bojonegoro Regency in 2021. A factor that causes the low quality of education in Bojonegoro district is that there has been no increase in equitable access to higher education. The scholarship program genius S1 Baznas Bojonegoro district is a tuition assistance program provided by the institution to students who are studying and domiciled in Bojonegoro district, where these students have a middle to lower economic level and have good academic potential. Classification methods that can be used are Random Forest and Multivariate Adaptive Regression Spline. There are 6 variable that influence scholarship acceptance, namely economic conditions, skills, performance, background, commitment and GPA. In this study, a Random Forest method produced an accuracy rate of 93.3% and an APER value of 6.7% with influential variables being skill variables, performance variables, GPA variables, commitment variables, background variables and economic condition variables. And MARS modeling is obtained by using trial and error using Basis Function (BF) 12, 18, 24 then the best model is obtained with Basis Function (BF) 18, Maximum Interaction value 2, Minimum Observation (MO) 0 and GCV value 0.07941 with an accuracy value of 91% and an APER value of 8.9% with influential variables are Skill, GPA, Performance, and Economic Conditions.

Keywords: *Education, Random Forest, MARS*

ABSTRAK

Achyani, N, I. (2023). *Random Forest Dan Multivariate Adaptive Regression Spline (MARS) Untuk Klasifikasi Penerimaan Beasiswa Genius S1 Baznas kabupaten Bojonegoro*. Skripsi. Prodi Statistika. Universitas Nahdlatul Ulama Sunan Giri Bojonegoro. Pembimbing Utama Alif Yuanita Kartini, M.Si. dan pembimbing kedua Fetrika Anggraini, M.Pd. Mutu pendidikan di Indonesia masih rendah. Di kabupaten Bojonegoro sendiri mutu pendidikan pada tahun 2021 masih dikatakan rendah hal ini dapat dilihat pada Angka Partisipasi Sekolah (APS) dan Angka Partisipasi Murni (APM). Faktor yang menyebabkan rendahnya mutu pendidikan di kabupaten Bojonegoro adalah karena belum adanya peningkatan pemerataan akses ke perguruan tinggi. Program beasiswa genius S1 Baznas kabupaten Bojonegoro adalah program bantuan biaya pendidikan yang diberikan kepada mahasiswa tersebut memiliki tingkat ekonomi yang kurang mampu. Metode klasifikasi yang dapat digunakan adalah Random Forest dan Multivariate Adaptive Regression Spline. Variabel-variabel yang mempengaruhi penerimaan beasiswa ada 6 yaitu kondisi ekonomi, skill, performance, latar belakang, komitmen dan IPK. Pada penelitian ini, metode Random Forest menghasilkan tingkat akurasi sebesar 93,3 % dan nilai APER sebesar 6,7% dengan variabel yang berpengaruh adalah variabel skill, variabel variabel Performance, variabel IPK, variabel komitmen, variabel latar belakang dan variabel kondisi ekonomi. Dan didapatkan pemodelan MARS dengan menggunakan cara trial and error menggunakan Fungsi Basis (BF) 12, 18, 24 maka didapatkan model terbaik yaitu dengan Fungsi Basis (BF) 18, nilai Maksimum Interaksi 2, Minimum Observasi (MO) 0 dan nilai GCV 0,07941 dengan nilai akurasi sebesar 91,1% dan nilai APER sebesar 8,9% dengan variabel yang berpengaruh adalah Skill, IPK, Performance, dan Kondisi Ekonomi.

Kata kunci: Pendidikan, RandomForest, MARS

DAFTAR ISI

	Halaman
PERYATAAN KEASLIAN TULISAN	III
HALAMAN PERSETUJUAN	IV
HALAMAN PENGESAHAN	V
HALAMAN MOTTO DAN PERSEMBAHAN	VI
KATA PENGANTAR	VII
ABSTRACT	VIII
ABSTRAK	X
DAFTAR ISI	XI
DAFTAR TABEL	XIII
DAFTAR GAMBAR	XIV
DAFTAR LAMPIRAN	XV
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian	5
1.4 Manfaat Penelitian	6
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI	7
2.1 Tinjauan Pustaka	7
2.2 Dasar Teori	11
2.2.1 Statistika Deskriptif.....	11
2.2.2 Random Forest	15
2.2.3 <i>Multivariate Addaptive Regression Spline (MARS)</i>	19
2.2.4 Ketepatan Klasifikasi	24
2.2.5 Beasiswa Baznas kabupaten Bojonegoro.....	25
BAB III METODE PENELITIAN	29
3.1 Sumber Data	29
3.2 Variabel Penelitian	29
3.3 Tahapan Analisis	31
BAB IV PEMBAHASAN	34
4.1 Statistika Deskriptif	34

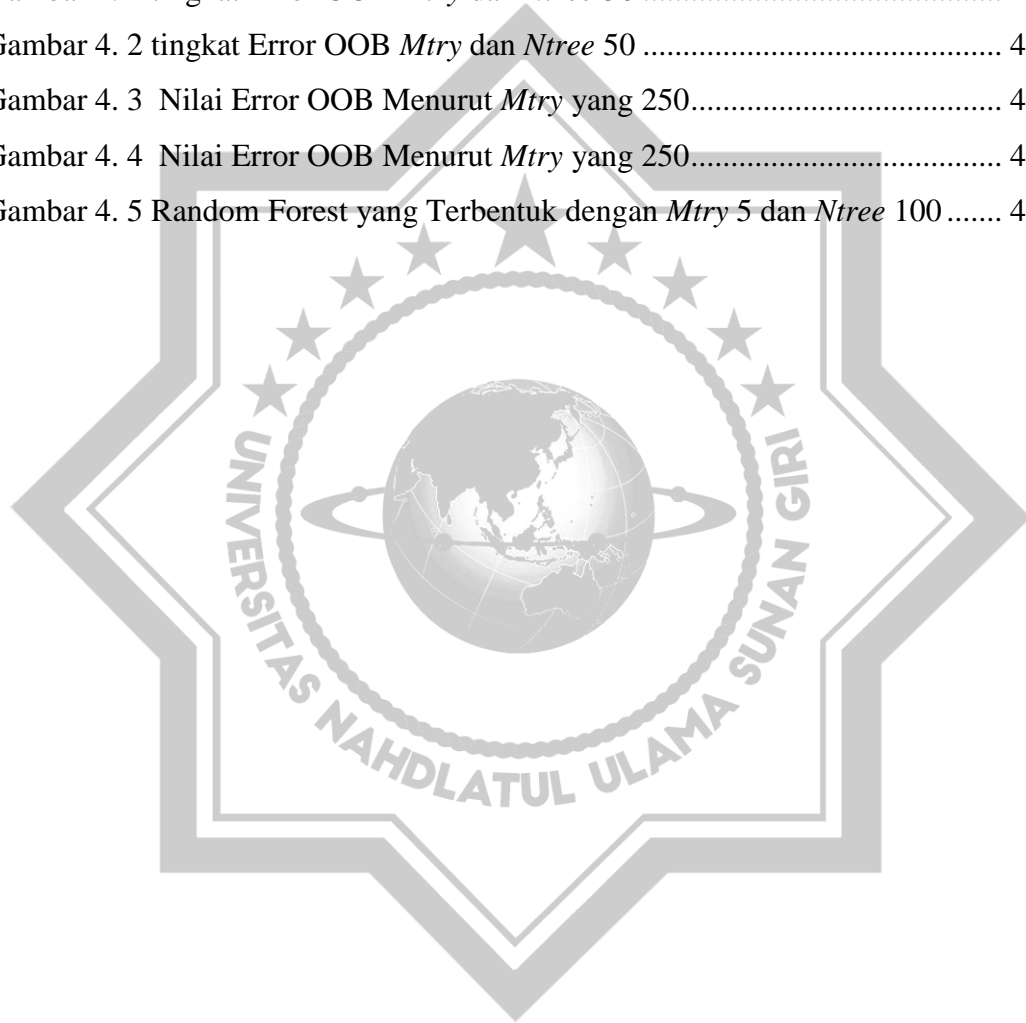
4.1.1	Status Calon Penerima Beasiswa Genius S1 Baznas	34
4.1.2	Kondisi Ekonomi	34
4.1.3	Skill	35
4.1.4	Performance	36
4.1.5	Latar Belakang	36
4.1.6	Komitmen.....	37
4.1.7	IPK (Indeks Prestasi Kumulatif)	37
4.2	Klasifikasi Random Forest.....	38
4.2.1	Penentuan Data Latih (<i>Training</i>) dan Data Uji (<i>Testing</i>).....	38
4.2.2	Penentuan <i>Mtry</i> dan Jumlah Pohon (<i>Ntree</i>) terbaik	38
4.2.3	Pendugaan Parameter Metode <i>Random Forest</i>	42
4.2.4	Pengujian Akurasi Metode <i>Random Forest</i>	44
4.3	Klasifikasi <i>Multivariate Adaptive Regression Spline</i> (MARS)	46
4.3.1	Penentuan Data Latih (<i>Training</i>) dan Data Uji (<i>Testing</i>).....	46
4.3.2	Permodelan MARS	46
4.3.3	Pendugaan Parameter Model MARS	50
4.3.4	Pengujian Signifikan Parameter Model MARS	51
4.3.5	Klasifikasi Model MARS.....	52
4.4	Perbandingan Klasifikasi Metode Random Forest dan Metode MARS.....	54
BAB V KESIMPULAN DAN SARAN		56
5.1	Kesimpulan.....	56
5.2	Saran	58
DAFTAR PUSTAKA		59
LAMPIRAN.....		64

DAFTAR TABEL

Tabel	Halaman
Tabel 2. 1 Penelitian Terdahulu	7
Tabel 2. 2 Tabel Klasifikasi Model MARS	24
Tabel 3. 1 Variabel Penelitian	29
Tabel 4. 1 Statistik Deskriptif Status.....	34
Tabel 4. 2 Statistika Deskriptif Kondisi Ekonomi	35
Tabel 4. 3 Statistika Deskriptif Skill	36
Tabel 4. 4 Statistika Deskriptif Performance	36
Tabel 4. 5 Statistika Deskriptif Latar Belakang	37
Tabel 4. 6 Statistika Deskriptif Komitmen	37
Tabel 4. 7 Statistika Deskriptif IPK	38
Tabel 4. 8 Data Latih (<i>Train</i>) dan Data Uji (<i>Test</i>)	38
Tabel 4. 9 . Nilai Error OOB <i>Mtry</i> dan <i>Ntree</i> 50	39
Tabel 4. 10 Nilai Error OOB <i>Mtry</i> dan <i>Ntree</i> 100	40
Tabel 4. 11 Nilai Error OOB <i>Mtry</i> dan <i>Ntree</i> 250	41
Tabel 4. 12 Nilai Error OOB <i>Mtry</i> dan <i>Ntree</i> 250	42
Tabel 4. 13 Tingkat Kepentingan Variabel dalam <i>Random Forest</i>	43
Tabel 4. 14 Hasil Prediksi Data Uji (<i>Testing</i>)	45
Tabel 4. 15 Data Latih (<i>Train</i>) dan Data Uji (<i>Test</i>)	46
Tabel 4. 16 Hasil Trial and Error Model MARS	47
Tabel 4. 17 Interaksi Antar Variabel Prediktor	49
Tabel 4. 18 Tingkat Kepentingan Variabel dalam MARS	50
Tabel 4. 19 Pengujian Signifikansi Model MARS	51
Tabel 4. 20 Klasifikasi Model MARS Menggunakan Data Uji(<i>Training</i>)	53
Tabel 4. 21 Nilai Akurasi Metode Random Forest dan MARS	54
Tabel 4. 22 Nilai APER Metode Random Forest dan MARS.....	55

DAFTAR GAMBAR

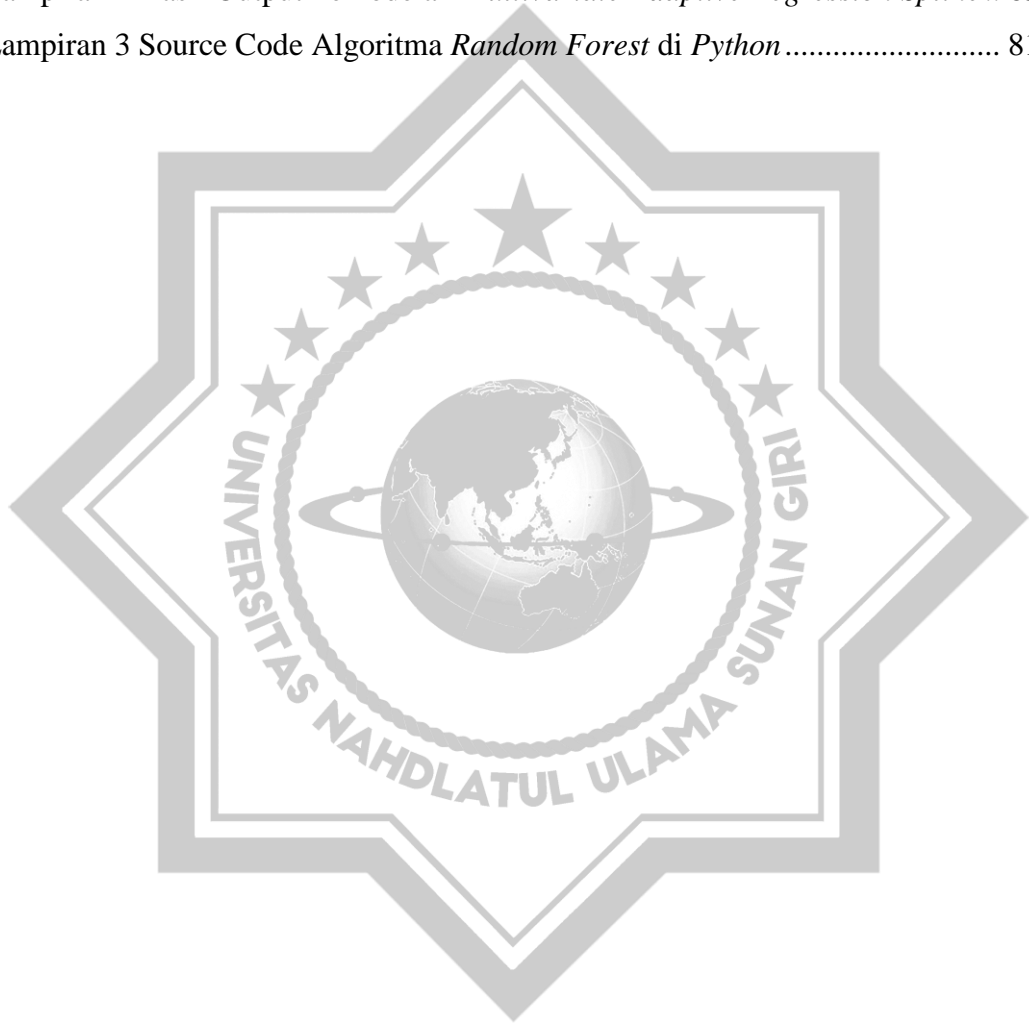
Gambar	Halaman
Gambar 2. 1 Proses Klasifikasi Random Forest.....	17
Gambar 3. 1 Diagram Alir MARS dan Random Forest.....	33
Gambar 4. 1 tingkat Error OOB <i>Mtry</i> dan <i>Ntree</i> 50	40
Gambar 4. 2 tingkat Error OOB <i>Mtry</i> dan <i>Ntree</i> 50	40
Gambar 4. 3 Nilai Error OOB Menurut <i>Mtry</i> yang 250.....	41
Gambar 4. 4 Nilai Error OOB Menurut <i>Mtry</i> yang 250.....	42
Gambar 4. 5 Random Forest yang Terbentuk dengan <i>Mtry</i> 5 dan <i>Ntree</i> 100	44



UNUGIRI

DAFTAR LAMPIRAN

Lampiran	Halaman
Lampiran 1 Data Calon Penerima Beasiswa Genius S1 Baznas Kabupaten Bojonegoro	64
Lampiran 2 Hasil Output Pemodelan <i>Multivariate Adaptive Regression Spline</i> ..	65
Lampiran 3 Source Code Algoritma <i>Random Forest</i> di <i>Python</i>	81



UNUGIRI