



**KLASIFIKASI PENERIMA PROGRAM KELUARGA
HARAPAN (PKH) MENGGUNAKAN METODE *RANDOM
FOREST* DI KELURAHAN 13 ULU, KECAMATAN
SEBERANG ULU II, PALEMBANG**

SKRIPSI

**AGUNG NASRULLAH
2020.11.0055**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER DAN SAINS
UNIVERSITAS INDO GLOBAL MANDIRI
2024**

**KLASIFIKASI PENERIMA PROGRAM KELUARGA
HARAPAN (PKH) MENGGUNAKAN METODE *RANDOM
FOREST* DI KELURAHAN 13 ULU, KECAMATAN
SEBERANG ULU II, PALEMBANG**



SKRIPSI

**Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada Program Studi
Teknik Informatika**

Oleh:

**AGUNG NASRULLAH
2020.11.0055**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER DAN SAINS
UNIVERSITAS INDO GLOBAL MANDIRI
2024**

LEMBAR PENGESAHAN SKRIPSI

**Klasifikasi Penerima Program Keluarga Harapan (PKH)
Menggunakan Metode *Random Forest* Di Kelurahan 13 Ulu,
Kecamatan Seberang Ulu II, Palembang**

Oleh

**Agung Nasrullah
NPM : 2020.11.0055**

Palembang, 25 Juli 2024

Pembimbing I



**Muhammad Haviz Irfani, S.Si., M.T.I
NIK : 2021.01.0291**

Pembimbing II



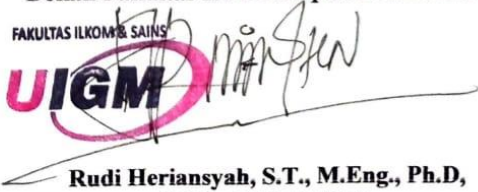
**Lastri Widya Astuti, M.Kom
NIK : 2003.01.0063**

Mengetahui,

Dekan Fakultas Ilmu Komputer dan Sains

FAKULTAS ILKOM & SAINS

UIGM


**Rudi Heriansyah, S.T., M.Eng., Ph.D,
NIK : 2022.01.0315**

LEMBAR PERSETUJUAN DEWAN PENGUJI

Pada hari Jum'at tanggal 5 Juli 2024 telah dilaksanakan ujian sidang skripsi :

Nama : Agung Nasrullah

NPM : 2020.11.0055

Judul : Klasifikasi Penerima Program Keluarga Harapan (PKH)
Menggunakan Metode *Random Forest* Di Kelurahan 13 Ulu,
Kecamatan Seberang Ulu II, Palembang

Oleh Prodi Teknik Informatika Fakultas Ilmu Komputer dan Sains Universitas
Indo Global Mandiri Palembang

Palembang, 23 Juli 2024

Penguji 1,



Rudi Heriansyah, S.T., M.Eng., Ph.D

NIK: 2022.01.0315

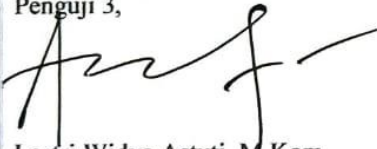
Penguji 2,



Dr. Shinta Puspasari, S.Si., M.Kom

NIK: 2015.01.0132

Penguji 3,



Lastri Widya Astuti, M.Kom

NIK: 2003.01.0063

Menyetujui,
Ka. Prodi Teknik Informatika



Zaid Romegar Mair, S.T., M.Cs

NIK: 2021.01.0307



SURAT KETERANGAN REVISI SKRIPSI
PROGRAM STUDI TEKNIK INFORMATIKA (S1)
FASILKOM DAN SAINS UNIVERSITAS INDO GLOBAL MANDIRI

Kami yang bertanda tangan dibawah ini, menerangkan bahwa :

Nama : Agung Nasrullah
NPM : 2020.11.0055
Judul : Klasifikasi Penerima Program Keluarga Harapan (PKH)
Menggunakan Metode Random Forest Di Kelurahan 13 Ulu,
Kecamatan Seberang Ulu II, Palembang

Mahasiswa yang namanya tercantum diatas, telah selesai merevisi penulisan SKRIPSI

Palembang, 23 Juli 2024

Penguji 1,

Rudi Heriansyah, S.T., M.Eng., Ph.D
NIK: 2022.01.0315

Penguji 2,

Dr. Shinta Puspasari, S.Si., M.Kom
NIK: 2015.01.0132

Penguji 3,

Lastris Widya Astuti, M.Kom
NIK: 2003.01.0063

Menyetujui,
Ka. Prodi Teknik Informatika

Zaid Romegar Mair, S.T., M.Cs
NIK: 2021.01.0307

**KLASIFIKASI PENERIMA PROGRAM KELUARGA
HARAPAN (PKH) MENGGUNAKAN METODE *RANDOM
FOREST* DI KELURAHAN 13 ULU, KECAMATAN
SEBERANG ULU II, PALEMBANG**

ABSTRAK

Program Keluarga Harapan (PKH) merupakan program bantuan sosial yang ditujukan bagi keluarga miskin dan rentan miskin yang terdaftar dalam Data Terpadu Kesejahteraan Sosial (DTKS) dengan persyaratan tertentu. Adapun penyaluran yang dilaksanakan sebelumnya belum tepat sasaran, dan pendataan masyarakat yang miskin masih belum terinci. Penelitian ini bertujuan melakukan klasifikasi penerima Program Keluarga Harapan (PKH) di Kelurahan 13 Ulu, Kecamatan Seberang Ulu II, Palembang. Metode yang digunakan yaitu Metode *Random Forest* dengan langkah *Preprocessing Data*, *Split Data*, Implementasi Metode *Random Forest*, *K-Fold Cross-Validation* dan Evaluasi. Atribut yang digunakan dalam penelitian meliputi, Kepemilikan Rumah, Jumlah Anak, Anak Sekolah, Anak Balita, Ibu Hamil, Lansia, Pekerjaan, Penghasilan, dan target atribut yaitu Kelayakan. Pada penelitian ini menggunakan 117 *dataset* yang sudah di *Preprocessing* dengan *Split Data* 70:30 menghasilkan *Accuracy* sebesar 97%, *Precision* sebesar 98%, *Recall* sebesar 87% dan *F1-Score* sebesar 92%. Kemudian dilanjutkan dengan *5 Fold Cross-Validation* menghasilkan rata-rata *Accuracy* sebesar 97%, *Precision* sebesar 95%, *Recall* sebesar 93% dan *F1-Score* sebesar 93%.

Kata Kunci: Program Keluarga Harapan (PKH), *Random Forest*, Klasifikasi Penerima Bantuan, Evaluasi Model, *K-Fold Cross-Validation*

***CLASSIFICATION OF RECIPIENTS OF THE PROGRAM
KELUARGA HARAPAN (PKH) USING THE RANDOM FOREST
METHOD IN KELURAHAN 13 ULU, SEBERANG ULU II SUB-
DISTRICT, PALEMBANG***

ABSTRACT

The Program Keluarga Harapan (PKH) is a social assistance program aimed at poor and vulnerable poor families registered in the Data Terpadu Kesejahteraan Sosial (DTKS) with certain requirements. The distribution that was carried out previously was not on target, and the data collection of poor people was still not detailed. This research aims to classify recipients of the Program Keluarga Harapan (PKH) in Kelurahan 13 Ulu, Seberang Ulu II District, Palembang. The method used is the Random Forest Method with the steps of Data Preprocessing, Split Data, Random Forest Method Implementation, K-Fold Cross-Validation and Evaluation. The attributes used in the research include, Home Ownership, Number of Children, School Children, Toddlers, Pregnant Women, Elderly, Employment, Income, and the target attribute is Eligibility. In this study using 117 datasets that have been preprocessed with Split Data 70:30 resulted in Accuracy of 97%, Precision of 98%, Recall of 87% and F1-Score of 92%. Then continued with 5 Fold Cross-Validation resulting in an average Accuracy of 97%, Precision of 95%, Recall of 93% and F1-Score of 93%.

Keywords: *Program Keluarga Harapan (PKH), Random Forest, Aid Recipient Classification, Model Evaluation, K-Fold Cross-Validation*

KATA PENGANTAR

Puji dan syukur penulis ucapkan kepada Allah SWT karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan penelitian pada skripsi ini tepat pada waktunya, tidak lupa juga shalawat beserta salam kepada Nabi Muhammad SAW beserta pengikutnya hingga pada akhir zaman.

Penelitian pada skripsi yang penulis buat dengan judul “Klasifikasi Penerima Program Keluarga Harapan (PKH) Menggunakan Metode *Random Forest* di Kelurahan 13 Ulu, Kecamatan Seberang Ulu II, Palembang” disusun guna memenuhi persyaratan untuk menyelesaikan Pendidikan Strata I Program Studi Teknik Informatika, Fakultas Ilmu Komputer dan Sains, Universitas Indo Global Mandiri Palembang.

Pertama dan paling utama, ucapan terima kasih kepada Allah SWT dan Nabi Muhammad SAW atas berkah dan karunia-Nyalah penulis bisa menyelesaikan penelitian pada skripsi ini. Selanjutnya, ucapan terima kasih penulis tujukan kepada yang terhormat:

1. Orangtua yang telah memberikan doa, restu serta dukungan yang sangat besar selama menjalani kehidupan hingga perkuliahan di Universitas Indo Global Mandiri.
2. Bapak Dr. H. Marzuki Alie, S.E., M.M, selaku Rektor Universitas Indo Global Mandiri.
3. Bapak Rudi Heriansyah, S.T., M.Eng., Ph.D, selaku Dekan Fakultas Ilmu Komputer dan Sains Universitas Indo Global Mandiri.
4. Bapak Zaid Romegar Mair, S.T., M.Cs, selaku Ketua Program Studi Teknik Informatika Universitas Indo Global Mandiri.
5. Bapak Muhammad Haviz Irfani, S.Si., M.T.I, selaku Dosen Pembimbing I yang telah membimbing saya selama pembuatan penelitian pada skripsi ini.
6. Ibu Latri Widya Astuti, M.Kom selaku Dosen Pembimbing II yang telah membimbing saya selama pembuatan penelitian pada skripsi ini.
7. Bapak Ir. Mustafa Ramadhan, M.T, sebagai Dosen Pembimbing Akademik.

8. Bapak/Ibu Dosen Program Studi Teknik Informatika yang telah memberikan banyak pembelajaran selama masa perkuliahan.
9. Teman-teman seperjuangan yang telah memberikan dukungan materi maupun moral sehingga penulis bisa menyelesaikan penelitian pada skripsi ini.

Penulis menyadari bahwa penelitian pada skripsi ini masih jauh dari kata sempurna, oleh karena itu kritik beserta saran untuk perbaikan dan pengembangan sangat dibutuhkan. Akhir kata, semoga penelitian pada skripsi ini bermanfaat bagi semua pihak. Penulis mengucapkan terima kasih.

Palembang, Juli 2024
Penulis,

Agung Nasrullah
2020.11.0055

DAFTAR ISI

HALAMAN JUDUL LUAR	i
HALAMAN JUDUL DALAM	ii
LEMBAR PENGESAHAN SKRIPSI	iii
LEMBAR PERSETUJUAN DEWAN PENGUJI	iv
SURAT KETERANGAN REVISI SKRIPSI	v
ABSTRAK	vi
ABSTRACT	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xvi
DAFTAR RUMUS	xvii
DAFTAR LAMPIRAN	xviii
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan dan Manfaat Penelitian.....	3
1.4.1 Tujuan Penelitian	3
1.4.2 Manfaat Penelitian	3
1.5 Sistematika Penulisan.....	4
BAB II LANDASAN TEORI	
2.1 Program Keluarga Harapan (PKH)	5
2.1.1 Tujuan Program Keluarga Harapan	5
2.1.2 Kriteria Penerima Manfaat Program Keluarga Harapan.....	6

2.1.3 Alur Pelaksanaan Program Keluarga Harapan	6
2.2 <i>Data Mining</i>	7
2.3 Klasifikasi.....	9
2.4 <i>Random Forest</i>	9
2.5 <i>Confusion Matrix</i>	12
2.6 <i>K-Fold Cross Validation</i>	13
2.7 <i>Python</i>	14
2.8 <i>Flowchart</i>	16
2.9 Penelitian Terdahulu.....	17
BAB III METODOLOGI PENELITIAN	
3.1 Tahapan Penelitian	21
3.2 Kebutuhan Perangkat	22
3.3 Observasi dan Wawancara	23
3.4 Studi Literatur.....	23
3.5 Identifikasi dan Perumusan Masalah.....	24
3.6 Pengumpulan Data	24
3.7 <i>Preprocessing Data</i>	26
3.8 Implementasi Metode <i>Random Forest</i>	28
3.9 Evaluasi Model.....	36
3.10 Pengambilan Kesimpulan.....	38
BAB IV HASIL DAN PEMBAHASAN	
4.1 Pengumpulan Data	39
4.2 Implementasi	39
4.2.1 <i>Data Loading</i>	39
4.2.2 <i>Preprocessing Data</i>	40

4.2.3 <i>Visualization Data</i>	46
4.2.4 <i>Split Data</i>	53
4.2.5 <i>Initialization Random Forest Classifier</i>	55
4.2.6 <i>Initialization K-Fold Cross-Validation</i>	56
4.2.7 <i>Evaluation</i>	57
BAB V PENUTUP	
5.1 Kesimpulan.....	68
5.2 Saran	68
DAFTAR PUSTAKA	69
LAMPIRAN	

DAFTAR GAMBAR

Gambar 2.1. Alur Pelaksanaan Program Keluarga Harapan.....	7
Gambar 2.2. Tahap <i>Data Mining</i>	8
Gambar 2.3. Ilustrasi Proses dari <i>Random Forest</i>	11
Gambar 2.4. Proses ilustrasi pada <i>K-Fold Cross Validation</i>	14
Gambar 3.1. Tahapan Penelitian	21
Gambar 3.2. Tahap <i>Preprocessing Data</i>	26
Gambar 3.3. Tabel Contoh Data Sampel Mentah	27
Gambar 3.4. Tahapan Implementasi <i>Random Forest</i>	29
Gambar 3.5. <i>Root</i> Pohon Keputusan	31
Gambar 3.6. <i>Node</i> 1.1 Pohon Keputusan	33
Gambar 3.7. <i>Node</i> 1.1.1 Pohon Keputusan	35
Gambar 3.8. Prediksi Pohon Keputusan pada Data Sampel	36
Gambar 4.1. <i>Loading Dataset</i>	39
Gambar 4.2. 10 Baris Pertama <i>Dataset</i>	40
Gambar 4.3. Normalisasi Data	41
Gambar 4.4. 10 Baris Pertama <i>Dataset</i> Setelah Dinormalisasi.....	42
Gambar 4.5. Kode Program untuk Menghapus Atribut Tidak Penting	42
Gambar 4.6. 10 Baris Pertama <i>Dataset</i> setelah Dihapus Atribut yang Tidak Penting	43
Gambar 4.7. Kode Program untuk Memeriksa Data Duplikat.....	43
Gambar 4.8. Jumlah Data Duplikat.....	44
Gambar 4.9. Kode Program untuk Menghapus Data Duplikat	44
Gambar 4.10. Kode Program untuk Memeriksa <i>Missing Value</i>	45
Gambar 4.11. Hasil Kode Program Memeriksa <i>Missing Value</i>	45
Gambar 4.12. Kode Program untuk Memeriksa Jumlah Data Setiap Atribut.....	46
Gambar 4.13. Jumlah Data Setiap Atribut	47
Gambar 4.14. Kode Program untuk Memeriksa Jumlah Data Label	47
Gambar 4.15. Jumlah Data Label.....	48

Gambar 4.16. Kode Program untuk Melihat Distribusi Data Setiap Atribut	48
Gambar 4.17. Distribusi Data Atribut ‘KEPEMILIKAN RUMAH’	49
Gambar 4.18. Distribusi Data Atribut ‘JUMLAH ANAK’	49
Gambar 4.19. Distribusi Data Atribut ‘ANAK SEKOLAH’	50
Gambar 4.20. Distribusi Data Atribut ‘BALITA’	50
Gambar 4.21. Distribusi Data Atribut ‘IBU HAMIL’	51
Gambar 4.22. Distribusi Data Atribut ‘LANSIA’	51
Gambar 4.23. Distribusi Data Atribut ‘PEKERJAAN’	52
Gambar 4.24. Distribusi Data Label ‘KELAYAKAN’	52
Gambar 4.25. Kode Program untuk Menentukan Fitur (X) dan Target (y).....	53
Gambar 4.26. Kode Program <i>Split Data</i>	53
Gambar 4.27. Kode Program untuk Melihat Jumlah <i>Data Train</i> dan <i>Test</i>	54
Gambar 4.28. Hasil Kode Program Melihat <i>Data Train</i> dan <i>Test</i>	54
Gambar 4.29. Kode Program untuk Inisialisasi <i>Random Forest Classifier</i>	55
Gambar 4.30. Inisialisasi <i>K-Fold Cross Validation</i>	56
Gambar 4.31. Kode Program <i>List</i>	57
Gambar 4.32. Kode Program untuk Menghitung <i>Recall</i> , <i>Precision</i> , Akurasi dan <i>F1-Score Testing</i>	58
Gambar 4.33. Hasil Kode Program Menghitung <i>Recall</i> , <i>Precision</i> , Akurasi dan <i>F1-Score Testing</i>	58
Gambar 4.34. Kode Program untuk Menghitung <i>Recall</i> , <i>Precision</i> , Akurasi dan <i>F1-Score Training</i>	59
Gambar 4.35. Hasil Kode Program Menghitung <i>Recall</i> , <i>Precision</i> , Akurasi dan <i>F1-Score Training</i>	59
Gambar 4.36. Kode Program untuk Mengukur <i>Confusion Matrix</i>	60
Gambar 4.37. <i>Confusion Matrix</i>	61
Gambar 4.38. Kode Program untuk Melakukan <i>K-Fold Cross Validation</i>	62
Gambar 4.39. Kode Program untuk Mengukur <i>Confusion Matrix K-Fold Cross Validation</i>	62
Gambar 4.40. Kode Program untuk Menghitung Rata-Rata di <i>K-Fold Cross Validation</i>	63

Gambar 4.41. Hasil Kode Program Mengukur *Confusion Matrix* dan Menghitung Rata-Rata pada *K-Fold Cross Validation*65

DAFTAR TABEL

Tabel 2.1. <i>Confusion Matrix</i>	13
Tabel 2.2. Simbol <i>Flowchart</i>	17
Tabel 2.3. Penelitian Terdahulu	17
Tabel 3.1. Atribut yang Digunakan.....	25
Tabel 3.2. Pembagian <i>Data Training</i> dan <i>Testing K-Fold Cross Validation</i>	28
Tabel 3.3. Perhitungan <i>Root</i> Pohon Keputusan	29
Tabel 3.4. Hasil Perhitungan <i>Root</i> Pohon Keputusan.....	30
Tabel 3.5. Perhitungan <i>Node 1.1</i> Pohon Keputusan	32
Tabel 3.6. Hasil Perhitungan <i>Node 1.1</i> Pohon Keputusan	32
Tabel 3.7. Perhitungan <i>Node 1.1.1</i> Pohon Keputusan.....	33
Tabel 3.8. Hasil Perhitungan <i>Node 1.1.1</i> Pohon Keputusan	34
Tabel 3.9. Perhitungan Data Sampel <i>Confusion Matrix</i>	37
Tabel 4.1. Tabel Perhitungan Rata-Rata <i>Recall</i> , <i>Precision</i> , Akurasi dan <i>F1-Score</i> Setiap <i>Fold</i>	66

DAFTAR RUMUS

Rumus 2.1. <i>Entropy Random Forest</i>	12
Rumus 2.2. <i>Information Gain Random Forest</i>	12
Rumus 2.3. <i>Recall Confusion Matrix</i>	13
Rumus 2.4. <i>Presisi Confusion Matrix</i>	13
Rumus 2.5. <i>Akurasi Confusion Matrix</i>	13
Rumus 2.6. <i>F1-Score Confusion Matrix</i>	13
Rumus 3.1. <i>Perhitungan Root Pohon Keputusan</i>	30
Rumus 3.2. <i>Perhitungan Node 1.1 Pohon Keputusan</i>	32
Rumus 3.3. <i>Perhitungan Node 1.1.1 Pohon Keputusan</i>	34
Rumus 3.4. <i>Perhitungan Data Sampel Recall Confusion Matrix</i>	38
Rumus 3.5. <i>Perhitungan Data Sampel Presisi Confusion Matrix</i>	38
Rumus 3.6. <i>Perhitungan Data Sampel Akurasi Confusion Matrix</i>	38
Rumus 3.7. <i>Perhitungan Data Sampel F1-Score Confusion Matrix</i>	38

DAFTAR LAMPIRAN

Lampiran 1 Daftar Riwayat Hidup

Lampiran 2 Kartu Bimbingan

Lampiran 3 Surat Pernyataan Tidak Plagiat

Lampiran 4 Surat Balasan Tempat Pengambilan Data Penelitian