

**KLASIFIKASI PENENTUAN PENERIMA PROGRAM  
KELUARGA HARAPAN (PKH) MENGGUNAKAN  
ALGORITMA *DECISION TREE* DI DESA TURIREJO,  
KABUPATEN GRESIK**

**SKRIPSI**

**Diajukan untuk memenuhi persyaratan dalam  
memperoleh gelar Sarjana Komputer Program**

**Studi Sistem Informasi**

**Disusun Oleh:**



**LEONA ELSA NILWANDA**

**19082010004**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"  
JAWA TIMUR S U R A B A Y A**

**2024**

**SKRIPSI**

**KLASIFIKASI PENENTUAN PENERIMA PROGRAM  
KELUARGA HARAPAN (PKH) MENGGUNAKAN  
ALGORITMA *DECISION TREE* DI DESA TURIREJO,  
KABUPATEN GRESIK**

Disusun Oleh:

**LEONA ELSA NILWANDA**

**19082010004**

Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Skripsi  
Program Studi Sistem Informasi Fakultas Ilmu Komputer  
Universitas Pembangunan Nasional "Veteran" Jawa Timur  
Pada Tanggal 8 Januari 2024

**Pembimbing:**

**Tim Penguji:**

1.

1.

**Amalia Anjani Arifiyanti, S.Kom., M.Kom**  
**NIP. 19920812 2018032 001**

**Eka Dyar Wahyuni, S.Kom., M.Kom.**  
**NIP. 19841201 2021212 005**

2.

2.

**Rizka Hadiwiyanti, S.Kom., M.Kom.**  
**NIP. 19860727 2018032 001**

**Abdul Rezha Efrat Najaf, S.Kom., M.Kom.**  
**NIP. 19940929 202203 1008**

3.

**Reisa Permatasari, S.T., M.Kom.**  
**NIP. 19920514 202203 2007**

**Mengetahui,**  
**Dekan Fakultas Ilmu Komputer**  
**Universitas Pembangunan Nasional "Veteran" Jawa Timur**

**Prof. Dr. Ir. Novirina Hendrasarie, M.T.**  
**NIP. 19681126 199403 2 001**



**LEMBAR PENGESAHAN**

**KLASIFIKASI PENENTUAN PENERIMA PROGRAM  
KELUARGA HARAPAN (PKH) MENGGUNAKAN  
ALGORITMA *DECISION TREE* DI DESA TURIREJO,  
KABUPATEN GRESIK**

**Disusun Oleh:**

**LEONA ELSA NILWANDA**  
**19082010004**

**Telah Disetujui mengikuti Ujian Negara Lisan Gelombang Januari  
Periode 2024 pada Tanggal 8 Januari 2024**

**Menyetujui,**

**Dosen Pembimbing 1**

**Dosen Pembimbing 2**

**Amalia Anjani Arjiyanti, S.Kom., M.Kom**  
**NIP. 19920812 2018032 001**

**Rizka Hadiwiyanti, S.Kom., M.Kom.**  
**NIP. 19860727 2018032 001**

**Mengetahui,**

**Ketua Program Studi Informasi**

**Fakultas Ilmu Komputer**

**Universtas Pembangunan Nasional "Veteran" Jawa Timur**

**Agung Brastama Putra, S. Kom., M.Kom**

**NIP. 19851124 2021211 003**





**KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"  
JAWA TIMUR  
FAKULTAS ILMU KOMPUTER**

**KETERANGAN REVISI**

Kami yang bertanda tangan di bawah ini menyatakan bahwa mahasiswa berikut:

Nama : Leona Elsa Nilwanda

NPM : 19082010004

Program Studi : Sistem Informasi

Telah mengerjakan revisi Ujian Negara Lisan Skripsi pada tanggal 8 Januari 2024 dengan judul :

**KLASIFIKASI PENENTUAN PENERIMA PROGRAM KELUARGA  
HARAPAN (PKH) MENGGUNAKAN ALGORITMA DECISION TREE DI  
DESA TURIREJO, KABUPATEN GRESIK**

Oleh karenanya mahasiswa tersebut diatas dinyatakan bebas revisi Ujian Negara Lisan Skripsi dan diijinkan untuk membukukan laporan SKRIPSI dengan judul tersebut.

Surabaya, Januari 2024

Dosen penguji yang memeriksa revisi:

1. Eka Dyar Wahyuni, S.Kom., M.Kom.

NIP. 19841201 2021212 005

2. Abdul Rezha Efrat Najaf, S.Kom., M.Kom.

NIP. 19940929 202203 1008

3. Reisa Permatasari, S.T., M.Kom.

NIP. 19920514 202203 2007

Mengetahui,

Dosen Pembimbing I

Amalia Anjani Arifiyanti, S.Kom., M.Kom.  
NIP. 19920812 2018032 001

Dosen Pembimbing II

Rizka Hadiwiyanti, S.Kom., M.Kom.  
NIP. 19860727 2018032 001





**KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"  
JAWA TIMUR  
FAKULTAS ILMU KOMPUTER**

**SURAT PERNYATAAN**

Saya yang bertanda tangan di bawah ini:

Nama : Leona Elsa Nilwanda  
NPM : 19082010004  
Program Studi : Sistem Informasi

Menyatakan bahwa Judul Skripsi / Tugas Akhir sebagai berikut:

**KLASIFIKASI PENENTUAN PENERIMA PROGRAM KELUARGA  
HARAPAN (PKH) MENGGUNAKAN ALGORITMA *DECISION TREE* DI  
DESA TURIREJO, KABUPATEN GRESIK**

Bukan merupakan plagiat dari Skripsi/Tugas Akhir / Penelitian orang lain dan juga bukan Produk / Hasil Karya yang saya beli dari orang lain.

Saya juga menyatakan bahwa Skripsi / Tugas Akhir ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional "Veteran" Jawa Timur maupun di Instansi Pendidikan lain. Jika dinyatakan dikemudian hari pernyataan terbukti benar, maka Saya bertanggung jawab penuh dan siap menerima segala konsekuensi, termasuk pembatalan ijazah dikemudian hari.

Surabaya, Januari 2024

Hormat Saya,



**Leona Elsa Nilwanda**  
**19082010004**

**Judul Skripsi** : **Klasifikasi Penerima Program Keluarga Harapan Menggunakan Algoritma *Decision tree* di Desa Turirejo, Kabupaten Gresik**

**Pembimbing 1** : **Amalia Anjani Arifiyanti, S.Kom.,M.Kom**

**Pembimbing 2** : **Rizka Hadiwiyanti, S.Kom.,M.Kom**

---

### ABSTRAK

Program Keluarga Harapan merupakan program bantuan sosial dari pemerintah untuk masyarakat miskin yang ditetapkan sebagai penerima Program Keluarga. Dengan adanya PKH ini masyarakat miskin di dorong untuk mendapat akses dan manfaat pelayanan sosial seperti kesehatan, pendidikan, pangan dan program perlindungan sosial lainnya.

Namun meskipun terbukti dapat menurunkan tingkat kemiskinan, dalam proses penyalurannya masih terjadi berbagai permasalahan, yang paling utama adalah penyaluran bantuan yang tidak tepat sasaran. Oleh karena itu, penelitian ini bertujuan untuk membangun sebuah model klasifikasi yang dapat digunakan untuk membantu proses klasifikasi penerima bantuan PKH. Algoritma yang digunakan dalam proses pemodelan adalah 3 algoritma *decision tree* yaitu ID3, C45 dan *Random forest* dengan bantuan *Feature Selection* PCA dan *Feature Importance* serta menggunakan SMOTE untuk mengatasi data yang tidak seimbang atau *Imbalance Data*. Dari hasil pemodelan kemudian dibandingkan dan yang memiliki hasil evaluasi paling baik akan di kembangkan dalam bentuk website.

Dari hasil pemodelan didapatkan bahwa model *Random forest* dengan SMOTE dan *Features Importance* memiliki *accuracy* yang paling tinggi di banding model yang lain sebesar 91%, *recall* 90%, *precision* 90%, *f1-score* 90% dan nilai ROC/AUC sebesar 93%. Hasil tersebut kemudian di uji validitasnya menggunakan 25 data baru pada sistem dengan hasil mendapat akurasi sebesar 92%.

**Kata Kunci:** Klasifikasi, *Decision tree*, PKH, *Random Forest*, ID3

## KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas Rahmat dan Karunia-Nya, penulis dapat menyelesaikan Penelitian Skripsi dengan judul “Klasifikasi Penentuan Penerima Program Keluarga Harapan (PKH) Menggunakan Algoritma *Decision Tree* Di Desa Turirejo, Kabupaten Gresik ” yang menjadi persyaratan dalam menyelesaikan Program Studi S1/Sistem Informasi di Universitas Pembangunan Nasional “Veteran” Jawa Timur ini dengan lancar.

Selama proses penyusunan penelitian skripsi ini, tidak lepas dari bimbingan, bantuan, serta masukan dari berbagai pihak, sehingga penulis pada kesempatan ini ingin mengucapkan terimakasih kepada:

1. Ibu Winarsih, Bapak Suwito, Kakak Siwid, Adik Vicka, beserta keluarga besar, atas doa, semangat, serta dukungannya kepada penulis dalam menyelesaikan penelitian skripsi.
2. Ibu Amalia Anjani Arifiyanti, S.Kom., M.Kom, selaku dosen pembimbing 1 yang telah memberikan ilmu, arahan, serta bimbingannya kepada penulis selama proses penelitian skripsi.
3. Ibu Rizka Hadiwiyanti, S.Kom., M.Kom., selaku dosen pembimbing 2 yang telah memberikan ilmu, arahan, serta bimbingannya kepada penulis selama proses penelitian skripsi.
4. Bapak Agung Brastama Putra, S.Kom., M.Kom., selaku koordinator Program Studi Sistem Informasi Universitas Pembangunan Nasional “Veteran” Jawa Timur yang telah membimbing penulis dalam mengerjakan penelitian skripsi.
5. Seluruh dosen Program Studi Sistem Informasi Universitas Pembangunan Nasional “Veteran” Jawa Timur yang telah memberikan ilmunya selama masa perkuliahan.
6. Seluruh pihak pemerintah desa Turirejo sebagai penyedia informasi yang telah membantu penulis selama masa penelitian.
7. Anisa dan Aisha sahabat yang telah memberikan dukungannya kepada penulis dan menemani penulis selama perkuliahan.
8. Seluruh teman-teman Mahasiswa Sistem Informasi Universitas Pembangunan Nasional “Veteran” Jawa Timur Angkatan 2019 penulis atas dukungannya.

9. Arctic Monkeys, Lana Del Rey, ONE OK ROCK, RINI, Amy Winehouse, Roy Orbison, DION, Alina Baraz yang telah menemani penulis dalam penyusunan skripsi dengan karya-karya musiknya.
10. Alex Turner, Wang Anyu, Eren Yeager, dan Levi Ackerman yang telah menjadi penyemangat bagi penulis dalam menyelesaikan skripsi.
11. Semua kucing-kucing peliharaan penulis, Kupus, Cici, Cemong, Selen, Gibro, Cemong II, Selin, Bocil, Cakil, Cemong III, Selon, Sappy, Gibro II yang telah menjadi penghibur bagi penulis saat lelah dalam penyusunan skripsi.

Dalam penulisan laporan skripsi ini, penulis menyadari bahwa penelitian skripsi ini masih jauh dari kata sempurna dan banyak kekurangannya. Oleh karena itu, penulis memohon maaf atas segala kekurangan yang ditemukan. Semoga penelitian skripsi ini memberikan manfaat, khususnya pada bidang sistem informasi.

Gresik, Januari 2024

Leona Elsa Nilwanda



## DAFTAR ISI

ABSTRAK .....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR .....	xiii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan.....	4
1.5 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Dasar Teori .....	6
2.1.1 PKH.....	6
2.1.2 Klasifikasi .....	8
2.1.3 <i>CRISP-DM</i> .....	8
2.1.4 <i>Random forest</i> .....	11
2.1.5 C4.5 .....	13
2.1.6 ID3.....	14
2.1.7 <i>Principal Component Analysis (PCA)</i> .....	15
2.1.8 <i>Confusion matrix</i> .....	16
2.1.9 <i>Classification Report</i> .....	18
2.1.9 ROC/AUC.....	19
2.1.10 AUC .....	20
2.1.11 <i>Features Importance</i> .....	21
2.1.12 <i>Flask</i> .....	22
2.1.13 SMOTE .....	22
2.2 Penelitian Terdahulu.....	23
2.2.1 Hasil Penelitian EKA Fitriani (2020).....	23
2.2.2 Penelitian Dwi Kinasih Widiyati dkk (2018).....	24
2.2.3 Penelitian Abdul Rofiq Almuqorobin (2021) .....	25
2.2.4 Penelitian Helen Sastyaprawatiwi dkk (2022).....	25

2.2.5	Penelitian Mia dkk (2022).....	26
BAB III METODOLOGI PENELITIAN.....		28
3.1	<i>Business Understanding</i> .....	29
3.1.1	Identifikasi Masalah.....	29
3.1.2	Studi Literatur .....	29
3.2	<i>Data Understanding</i> .....	29
3.2.1	Data Requirements .....	29
3.2.2	Data Collection.....	30
3.3	<i>Data Preparation</i> .....	1
3.4	<i>Modelling</i> .....	32
3.5	<i>Evaluation</i> .....	32
3.6	<i>Deployment</i> .....	33
3.6.1	Kebutuhan Pengembangan Sistem.....	33
BAB IV HASIL DAN PEMBAHASAN .....		35
4.1	Pengumpulan Data.....	35
4.2	<i>Data Preparation</i> .....	37
4.2.1	<i>Data Exploratory</i> .....	37
4.2.2	<i>Preprocessing</i> .....	39
4.2.3	<i>Visualization Data</i> .....	40
4.2.4	<i>Data Integration and Transformation</i> .....	51
4.3	<i>Modelling and Evaluation</i> .....	51
4.3.1	Skenario 1: <i>Random forest, ID3, C45</i> .....	52
4.3.2	Skenario 2: <i>Random forest, ID3, C45 + PCA</i> .....	54
4.3.3	Skenario 3: <i>Random forest, ID3, C45 + SMOTE</i> .....	58
4.3.4	Skenario 4: <i>Random forest, ID3, C45 + SMOTE+PCA</i> .....	61
4.3.5	Skenario 5: <i>Random forest, ID3, C45 + SMOTE + Features Importance</i> .....	64
4.3.6	Hasil Perbandingan Skenario .....	67
4.4	Pengembangan Sistem.....	68
4.5	Uji Validitas Sistem.....	70
BAB V PENUTUP.....		72
5.1	Kesimpulan .....	72
5.2	Saran.....	73
DAFTAR PUSTAKA .....		74

LAMPIRAN 1 .....	76
LAMPIRAN 2 .....	84



## DAFTAR TABEL

Tabel 2. 1 Nominal Bantuan PKH (Kementrian Sosial, 2019).....	7
Tabel 2. 2 Nilai AUC (Gorunescu, 2011) .....	21
Tabel 3. 1 Atribut Data Registrasi Sosial Ekonomi .....	30
Tabel 3. 2 Matriks Kebutuhan Fungsional.....	33
Tabel 4. 1 Tabel <i>Splitting</i> Data .....	52
Tabel 4. 2 Skenario Pertama <i>Random forest</i> , ID3, C4.5 .....	53
Tabel 4. 3 Skenario <i>Random forest</i> , ID3, C45 + PCA.....	56
Tabel 4. 4 Skenario <i>Random forest</i> , ID3, C45 + SMOTE .....	59
Tabel 4. 5 Skenario <i>Random forest</i> , ID3, C45 + SMOTE + PCA .....	62
Tabel 4. 6 Skenario <i>Random Forest</i> , ID3, C45 + SMOTE + <i>Features Importance</i> .....	65
Tabel 4. 7 Hasil Perbandingan Skenario .....	67
Tabel 4. 8 Tabel <i>Confussion Matrix</i> Uji Validitas Sistem .....	70

## DAFTAR GAMBAR

Gambar 2. 1 Tahapan <i>CRISP-DM</i> (Kenneth Jensen, 2016).....	9
Gambar 2. 2 <i>Random forest</i> (Hyacinth Ampadu, 2021) .....	11
Gambar 2. 3 <i>Confusion matrix</i> (Anuganti Suresh, 2020).....	17
Gambar 2. 4 Kurva ROC (Reza Faisal, 2017) .....	20
Gambar 2. 5 AUC (Google Developers, 2022).....	21
Gambar 3. 1 Metodologi Penelitian .....	28
Gambar 3. 3 Usecase Diagram.....	34
Gambar 4. 1 Informasi Dataset .....	38
Gambar 4. 2 <i>Missing Value</i> .....	38
Gambar 4. 3 <i>Columns Single Value Code</i> .....	39
Gambar 4. 4 <i>Cleansing Data Using Mean and Modus</i> .....	39
Gambar 4. 5 Bantuan.....	40
Gambar 4. 6 Rumah .....	41
Gambar 4. 7 Statistik Penerima Bantuan Berdasarkan Rumah.....	41
Gambar 4. 8 Luas .....	42
Gambar 4. 9 Statistik Penerima Bantuan Berdasarkan Luas.....	42
Gambar 4. 10 Lantai.....	43
Gambar 4. 11 Statistik Penerima Bantuan Berdasarkan Lantai .....	43
Gambar 4. 12 Dinding.....	44
Gambar 4. 13 Statistik Penerima Bantuan Berdasarkan Dinding .....	44
Gambar 4. 14 Atap .....	45
Gambar 4. 15 Statistik Penerima Bantuan Berdasarkan Atap.....	45
Gambar 4. 16 Air.....	46

Gambar 4. 17 Statistik Penerima Bantuan Berdasarkan Air .....	46
Gambar 4. 18 Listrik .....	47
Gambar 4. 19 Statistik Penerima Bantuan Berdasarkan Listrik.....	47
Gambar 4. 20 Umur.....	48
Gambar 4. 21 Pendidikan.....	48
Gambar 4. 22 Statistik Penerima Bantuan Berdasarkan Pendidikan .....	49
Gambar 4. 23 Tanggungan.....	49
Gambar 4. 24 Statistik Penerima Bantuan Berdasarkan Tanggungan .....	50
Gambar 4. 25 Statistik Penerima Bantuan Berdasarkan Pekerjaan.....	50
Gambar 4. 26 <i>Source code</i> proses Encoding.....	51
Gambar 4. 27 Alur Pemodelan Skenario <i>Random forest</i> , ID3, C45 .....	52
Gambar 4. 28 <i>Confusion Matrix Ranfom Forest</i> .....	53
Gambar 4. 29 <i>Classification Report Random Forest</i> .....	54
Gambar 4. 30 Alur Pemodelan Skenario <i>Random forest</i> , ID3, C45 + PCA .....	55
Gambar 4. 31 <i>Confusion Matrix ID3 + 7 PCs</i> .....	57
Gambar 4. 32 <i>Classification Report ID3 + 7 PCs</i> .....	58
Gambar 4. 33 Alur Pemodelan Skenario <i>Random forest</i> , ID3, C45 + SMOTE ...	59
Gambar 4. 34 <i>Confusion Matrix Random Forest + SMOTE</i> .....	60
Gambar 4. 35 <i>Classification Report Random Forest + SMOTE</i> .....	60
Gambar 4. 36 Alur Pemodelan Skenario <i>Random forest</i> , ID3, C45 + SMOTE + PCA.....	61
Gambar 4. 37 <i>Confusion Matrix Random Forest + SMOTE + 10 PCs</i> .....	63
Gambar 4. 38 <i>Classification Report Random Forest + SMOTE + 10 PCs</i> .....	64



Gambar 4. 39 Skenario Random Forest, ID3, C45 + SMOTE + Features	
Importance.....	65
Gambar 4. 40 <i>Confusion Matrix Random Forest + SMOTE + Features</i>	
<i>Importance</i> .....	66
Gambar 4. 41 <i>Classification Report Random Forest + SMOTE + Features</i>	
<i>Importance</i> .....	67
Gambar 4. 45 Beranda Aplikasi PKH.....	68
Gambar 4. 46 Halaman Prediksi .....	69
Gambar 4. 47 Halaman Visualisasi.....	69