

**PENERAPAN ALGORITMA *ITERATIVE DICHOTOMISER*
THREE (ID3) UNTUK MENENTUKAN PENERIMA
PROGRAM HIBAH AIR MINUM PERDESAAN DI
KABUPATEN JEPARA**



SKRIPSI

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Strata 1
(S.1) Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Nahdlatul Ulama Jepara

Oleh:

Bunga Alfiena Darissalamah

NIM: 171240000722

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NAHDLATUL ULAMA JEPARA
2021**

**PENERAPAN ALGORITMA *ITERATIVE DICHOTOMISER*
THREE (ID3) UNTUK MENENTUKAN PENERIMA
PROGRAM HIBAH AIR MINUM PERDESAAN DI
KABUPATEN JEPARA**



SKRIPSI

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Strata 1
(S.1) Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Nahdlatul Ulama Jepara

Oleh:

Bunga Alfiena Darissalamah

NIM: 171240000722

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NAHDLATUL ULAMA JEPARA
2021**

PERSETUJUAN PEMBIMBING

Assalamu'alaikum Wr. Wb.

Setelah kami meneliti dan mengadakan perbaikan seperlunya, bersama ini saya kirim naskah Skripsi Saudara:

Nama : Bunga Alfiena Darissalamah
NIM : 171240000722
Program Studi : Teknik Informatika
Judul : Penerapan Algoritma *Iterative Dichotomiser Three* (ID3) untuk Menentukan Penerima Program Hibah Air Minum Perdesaan di Kabupaten Jepara

Skripsi ini telah disetujui dan siap dipertahankan dihadapan tim penguji program Sarjana Strata 1 (S1) Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama (UNISNU) Jepara.

Demikian harap menjadikan maklum.

Wassalamu'alaikum Wr. Wb.

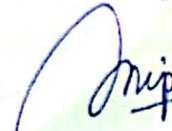
Jepara, 12 April 2021

Pembimbing I



Harminto Mulyo, S.Kom., M.Kom.
NIDN. 0604028203

Pembimbing II



Akhmad Khanif Zyen, M.Kom.
NIDN. 0621048602

Mengetahui,
Kepala Program Studi Teknik Informatika



Akhmad Khanif Zyen, M.Kom.
NIDN. 0621048602

PENGESAHAN

Skripsi dengan judul “Penerapan Algoritma *Iterative Dichotomiser Three* (ID3) untuk Menentukan Penerima Program Hibah Air Minum Perdesaan di Kabupaten Jepara” karya:

Nama : Bunga Alfiena Darissalamah
NIM : 171240000722
Program Studi : Teknik Informatika

Telah diujikan dan dipertahankan dalam sidang oleh Dewan Penguji Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama (UNISNU) Jepara dan dinyatakan lulus pada tanggal : 27 April 2021

Selanjutnya dapat diterima sebagai syarat guna memperoleh gelar Sarjana Strata 1 (S1) Program Studi Teknik Informatika pada Fakultas Sains dan Teknologi UNISNU Jepara Tahun Akademik 2021.

Jepara, 27 April 2021

Ketua Sidang

Sekretaris Sidang



Harminto Mulyo, S.Kom., M.Kom.
NIDN. 0604028203

Penguji I



Akhmad Khanif Zyen, M.Kom.
NIDN. 0621048602

Penguji II



Sarwido, S.E., M.M.
NIDN. 0625016902



Nur Aeni Widiastuti, S.Pd., M.Kom.
NIDN. 0602078702

Mengetahui,
Dekan Fakultas Sains dan Teknologi
UNISNU Jepara



Ir. Gun Sudiryanto, M.M.
NIDN. 0624056501

PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama : Bunga Alfiena Darissalamah

NIM : 171240000722

Program Studi : Teknik Informatika

Saya menyatakan dengan penuh kejujuran dan tanggungjawab, bahwa Skripsi yang saya susun sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata 1 (S1) Universitas Islam Nahdlatul Ulama Jepara merupakan hasil karya sendiri yang jauh dari plagiarisme dan belum pernah diajukan sebagai pemenuhan persyaratan untuk memperoleh gelar Sarjana dari Perguruan Tinggi lain.

Adapun bagian-bagian tertentu dalam penulisan Skripsi yang saya kutip dari karya orang lain telah dituliskan sumbernya secara jelas dengan norma, kaidah, dan etika penulisan ilmiah.

Selanjutnya saya bersedia menerima sanksi dari Fakultas Sains dan Teknologi UNISNU Jepara apabila dikemudian hari ditemukan ketidakbenaran dari pernyataan ini.

Jepara, 12 April 2021

Peneliti,



Bunga Alfiena Darissalamah

NIM. 171240000722

ABSTRAK

Judul	: Penerapan Algoritma <i>Iterative Dichotomiser Three</i> (ID3) untuk Menentukan Penerima Program Hibah Air Minum Perdesaan di Kabupaten Jepara
Penulis	: Bunga Alfiena Darissalamah
NIM	: 171240000722
Program Studi	: Teknik Informatika
Pembimbing I	: Harminto Mulyo, S.Kom., M.Kom.
Pembimbing II	: Akhmad Khanif Zyen, M.Kom.
Penguji I	: Sarwido, M.M.
Penguji II	: Nur Aeni Widiastuti, S.Pd., M.Kom.
Tanggal Ujian	: 27 April 2021

Program Hibah Air Minum Perdesaan (HAMP) merupakan sebuah program yang dibuat oleh pemerintah dalam upaya pemenuhan akses air minum dengan memberikan bantuan Sambungan Rumah (SR) kepada masyarakat. Namun, dalam proses penyeleksiannya masih terdapat kendala untuk menentukan penerima, karena banyaknya jumlah data yang harus diperiksa satu-persatu sehingga memakan waktu yang lama dan belum tepat sasaran dalam menentukan penerima. Dari permasalahan tersebut, dibutuhkan suatu metode untuk menentukan penerima yang tepat sasaran dan efektif. Data yang digunakan dalam penelitian ini berjumlah 1050 *record* dengan 15 atribut yaitu id calon penerima, kecamatan, desa/kelurahan, nama, alamat (RT/RW), nama responden, kesesuaian alamat, hubungan keluarga, jenis kelamin, daya listrik, status sambungan PDAM, kesiediaan menyambung, hasil survey, alasan *ineligible*, dan lainnya, yang akan diolah menggunakan Algoritma *Iterative Dichotomiser Three* (ID3). Kemudian dilakukan pengujian model menggunakan *K-Fold Cross Validation* dan *Confusion Matrix* untuk evaluasi dan validasi hasil yang menghasilkan akurasi sebesar 98,73% dengan nilai AUC (*Area Under Curve*) sebesar 0,981, sehingga dapat dinilai bahwa Algoritma *Iterative Dichotomiser Three* (ID3) mampu bekerja dengan sangat baik. Dengan kelebihan yang dimiliki Algoritma *Iterative Dichotomiser Three* (ID3), maka pada penelitian ini dihasilkan pohon keputusan dengan 10 *rules* yang diharapkan dapat memudahkan dalam menentukan penerima Program Hibah Air Minum Perdesaan (HAMP) di Kabupaten Jepara.

Kata kunci: Program Hibah Air Minum Perdesaan (HAMP), Algoritma *Iterative Dichotomiser Three* (ID3), *K-Fold Cross Validation*, *Confusion Matrix*, *Area Under Curve*.

ABSTRACT

Title : *Application of the Iterative Dichotomiser Three (ID3) Algorithm to Determine Recipients of the Rural Drinking Water Grant Program in Jepara Regency*
Author : *Bunga Alfiena Darissalamah*
NIM : *171240000722*
Study Program : *Informatics Engineering*
Supervisor I : *Harminto Mulyo, S.Kom., M.Kom.*
Supervisor II : *Akhmad Khanif Zyen, M.Kom.*
Tester I : *Sarwido, M.M.*
Tester II : *Nur Aeni Widiastuti, S.Pd., M.Kom.*
Test Date : *27 April 2021*

The Rural Drinking Water Grant Program (HAMP) is a program created by the government in an effort to fulfill access to drinking water by providing house connection (SR) assistance to the community. However, in the selection process there were still obstacles in determining the recipient, because the large amount of data had to be checked one by one, so it took a long time and was not on target in determining the recipient. From these problems, we need a method to determine the right target and effective recipients. The data used in this study amounted to 1050 records with 15 attributes, namely the ID of the prospective recipient, sub-district, village / sub-district, name, address (RT / RW), name of the respondent, suitability of address, family relationship, gender, electrical power, PDAM connection status. , willingness to connect, survey results, ineligible reasons, and others, which will be processed using the Iterative Dichotomiser Three (ID3) Algorithm. Then the model testing was carried out using the K-Fold Cross Validation and Confusion Matrix for evaluation and validation of the results which resulted in an accuracy of 98.73% with an AUC (Area Under Curve) value of 0.981, so that it can be judged that the Iterative Dichotomiser Three (ID3) Algorithm is able to work. very well. With the advantages possessed by the Iterative Dichotomiser Three (ID3) Algorithm, this research produces a decision tree with 10 rules which is expected to make it easier to determine the recipient of the Rural Drinking Water Grant Program (HAMP) in Jepara Regency.

Keywords: *Rural Drinking Water Grant Program (HAMP), Algorithm Iterative Dichotomiser Three (ID3), K-Fold Cross Validation, Confusion Matrix, Area Under Curve.*

MOTTO

“Allah tidak akan membebani seseorang melainkan sesuai dengan kadar kesanggupannya.”

(Q.S Al-Baqarah : 286)

“Kalo lo gagal SNMPTN, SBMPTN, Mandiri, gagal dapet negeri, itu biasa. Mungkin kedengarannya jahat, tapi gw pernah ngalamin itu semua. Dan itu bukan akhirnya, itu justru awal dari semua hal besar. Gapapa, semua orang pasti akan sampai tujuan. Kendaraan dan waktunya aja yang beda.”

(Rintik Sedu)

“Jika ingin mendapatkan yang lebih. Maka, usahanya juga harus lebih.”

(Author | Ganteng Sumpah)

PERSEMBAHAN

Skripsi ini saya persembahkan kepada:

1. Bpk. Mahsun dan Ibu Himmatul Hasanah, kedua orang tua peneliti. Sebab selalu mendo'akan apa-apa saja kegiatan yang dilakukan peneliti, memberikan nasihat-nasihat yang membangun semangat, hingga menjadi sebuah motivasi bagi peneliti dalam menempuh pendidikan di bangku perkuliahan hingga menyelesaikan Skripsi ini.
2. Abdul Hamid dan Renny Hastutik, dua saudara peneliti yang selalu ada untuk membantu peneliti.
3. Teman-teman peneliti di perkuliahan, TIF17, HMPSIF, BEM FST, dan KKN Kelompok 24, serta teman peneliti selama di SMA Patra Mandiri 01 Palembang yang telah membantu dan melengkapi hari peneliti dengan kesenangan.
4. Park Chanyeol, yang telah memberikan semangat melalui karya dan memotivasi peneliti untuk terus disiplin dan berambisi dalam mencapai keinginan yang diharapkan.

KATA PENGANTAR

Segala puji dan syukur kehadirat Allah SWT yang telah memberikan rahmat, taufiq, dan hidayah-Nya, sehingga peneliti dapat menyelesaikan penelitian dengan judul: “Penerapan Algoritma *Iterative Dichotomiser Three* (ID3) untuk Menentukan Penerima Program Hibah Air Minum Perdesaan di Kabupaten Jepara” dengan baik.

Peneliti menyadari dalam menyelesaikan Skripsi ini, terdapat keterbatasan dari pengetahuan yang dimiliki. Oleh karena itu, tanpa keterlibatan dan sumbangsih dari berbagai pihak, sulit bagi peneliti untuk menyelesaikan penyusunan Skripsi ini. Maka, dengan segala kerendahan hati serta rasa syukur, peneliti mengucapkan terima kasih sebesar-besarnya kepada:

1. Rektor Universitas Islam Nahdlatul Ulama (UNISNU) Jepara, Bpk. Dr. H. Sa’dullah Assaidi, M.Ag., yang telah memberikan ilmu pengetahuan sehingga dapat menjadi rasa semangat bagi peneliti dalam menempuh pendidikan.
2. Dekan Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama (UNISNU) Jepara, Bpk. Ir. Gun Sudiryanto, M.M., yang telah memfasilitasi dan memberikan kemudahan bagi peneliti sehingga dapat menyelesaikan Skripsi dengan baik.
3. Ketua Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama (UNISNU) Jepara, Bpk. Akhmad Khanif Zyen, M.Kom., yang telah memberikan arahan dan saran kepada peneliti selama menempuh pendidikan.
4. Pembimbing I, Bpk. Harminto Mulyo, S.Kom., M.Kom. dan pembimbing II, Bpk. Akhmad Khanif Zyen, M.Kom., yang telah membimbing peneliti dengan penuh kesabaran dan tanggungjawab dalam penyusunan skripsi hingga selesai.
5. Bapak dan Ibu Dosen Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama (UNISNU) Jepara, yang telah

memberikan ilmu pengetahuan serta wawasannya kepada peneliti selama perkuliahan.

6. Pihak Dinas Perumahan Rakyat dan Kawasan Permukiman (DISPERKIM) Kabupaten Jepara, yang telah memberikan izin kepada peneliti untuk melaksanakan penelitian hingga selesai.
7. Teman angkatan tahun 2017 Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama (UNISNU) Jepara, yang telah membantu dan memberikan semangat kepada peneliti.

Peneliti juga menyadari bahwa dalam penulisan Skripsi ini masih jauh dari kata sempurna. Maka, kritik dan saran yang bersifat membangun sangat dibutuhkan guna menyempurnakan Skripsi ini nantinya. Peneliti berharap Skripsi ini dapat memberikan manfaat bagi peneliti sendiri maupun pembaca yang melihat dan membacanya.

Jepara, 12 April 2021

Peneliti,

Bunga Alfiena Darissalamah

NIM. 171240000722

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN PEMBIMBING	ii
HALAMAN PENGESAHAN	iii
PERNYATAAN KEASLIAN	iv
HALAMAN ABSTRAK	v
HALAMAN MOTTO	vii
HALAMAN PERSEMBAHAN	viii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian.....	6
1.6 Sistematika Penulisan.....	7
BAB II LANDASAN TEORI	8
2.1 Tinjauan Studi	8
2.2 Tinjauan Pustaka	10
2.2.1 Program Hibah Air Minum Perdesaan (HAMP).....	10
2.2.2 Data Mining	11

2.2.3	Algoritma Iterative Dichotomiser Three (ID3)	13
2.2.4	Pohon Keputusan	26
2.2.5	<i>Cross Validation</i>	28
2.2.6	<i>Confusion Matrix</i>	28
2.2.7	Kurva ROC (<i>Receiver Operating Characteristic</i>)	29
2.2.8	<i>RapidMiner</i>	30
2.3	Kerangka Pemikiran	32
BAB III METODOLOGI PENELITIAN		33
3.1	Desain Penelitian	33
3.2	Pengumpulan Data	34
3.3	Lokasi Penelitian	35
3.4	Pengolahan Data Awal	35
3.5	Metode yang Diusulkan.....	37
3.6	Pengujian Model.....	37
3.7	Evaluasi dan Validasi Hasil.....	37
BAB IV HASIL DAN PEMBAHASAN		38
4.1	Pengumpulan Data	38
4.2	Pengolahan Data Awal	38
4.2.1	Menyusun dan Mengkategorikan Data	39
4.2.2	Seleksi Atribut.....	41
4.2.3	Pembersihan Data.....	44
4.3	Pengujian Model <i>K-Fold Cross Validation</i>	45
4.4	Algoritma <i>Iterative Dichotomiser Three (ID3)</i>	50
4.4.1	Menghitung Jumlah Sampel.....	52
4.4.2	Menghitung <i>Entropy</i> (Total) untuk <i>Root Node</i>	52

4.4.3	Menghitung <i>Entropy Ratio</i> Tiap Atribut untuk <i>Root Node</i>	52
4.4.4	Menghitung <i>Information Gain</i> Tiap Atribut dan Menentukan Nilai Tertinggi untuk <i>Root Node</i>	57
4.4.5	Menentukan Cabang dari Tiap-Tiap Atribut.....	60
4.5	Pohon Keputusan dan Aturan.....	91
4.6	Evaluasi dan Validasi Hasil.....	93
4.6.1	<i>Confusion Matrix</i>	93
4.6.2	Kurva ROC (<i>Receiver Operating Characteristic</i>)	96
BAB V KESIMPULAN		97
5.1	Kesimpulan.....	97
5.2	Saran	97
DAFTAR PUSTAKA		98
LAMPIRAN-LAMPIRAN		98

DAFTAR GAMBAR

Gambar 2.1 Proses <i>Knowlegde Discovery in Databases</i> (KDD).....	11
Gambar 2.2 Flowchart Algoritma <i>Iterative Dichotomiser Three</i> (ID3).....	15
Gambar 2.3 Pohon Keputusan	27
Gambar 2.4 Tampilan <i>RapidMiner</i>	31
Gambar 2.5 Kerangka Pemikiran.....	32
Gambar 3.1 <i>Flowchart</i> Penelitian	33
Gambar 4.1 Hasil Penyusunan dan Pengkategorian Data.....	41
Gambar 4.2 Permodelan Seleksi Atribut.....	42
Gambar 4.3 <i>Parameters Select Attributes</i>	42
Gambar 4.4 Seleksi Atribut Gambar 4. 5.....	43
Gambar 4.6 Hasil Seleksi Atribut	43
Gambar 4.7 Permodelan <i>Replace Missing Values</i>	44
Gambar 4.8 <i>Parameters Replace Missing Values</i>	45
Gambar 4.9 Hasil Pembersihan Data	45
Gambar 4.10 Permodelan <i>K-Fold Cross Validation</i>	46
Gambar 4.11 Permodelan Algoritma <i>Iterative Dichotomiser Three</i> (ID3).....	46
Gambar 4.12 Permodelan Split Data.....	47
Gambar 4.13 <i>Parameters Split Data</i>	47
Gambar 4.14 Hasil Pengujian Model Pertama.....	48
Gambar 4.15 Hasil Pengujian Model Kedua.....	48
Gambar 4.16 Hasil Pengujian Model Ketiga	49
Gambar 4.17 Hasil Pengujian Model Keempat.....	49
Gambar 4.18 Hasil Pengujian Model Kelima	50
Gambar 4.19 Pohon Keputusan.....	92
Gambar 4.20 <i>Performance Accuracy</i>	94
Gambar 4.21 <i>Performance Precision</i>	95
Gambar 4.22 <i>Performance Recall</i>	95
Gambar 4.23 Kurva ROC (<i>Receiver Operating Characteristic</i>)	96

DAFTAR TABEL

Tabel 1.1 Kriteria Penerima Program Hibah Air Minum Perdesaan (HAMP).....	3
Tabel 2.1 Contoh <i>Dataset</i>	16
Tabel 2.2 Hasil Perhitungan <i>Gain</i> Tertinggi pada <i>Node</i> 1	20
Tabel 2.3 Hasil Perhitungan <i>Gain</i> Tertinggi pada <i>Node</i> 1.1.....	23
Tabel 2.4 Hasil Perhitungan <i>Gain</i> Tertinggi pada <i>Node</i> 1.1.2.....	26
Tabel 2.5 Contoh Pembagian <i>Dataset</i>	28
Tabel 2.6 <i>Confusion Matrix</i>	28
Tabel 3.1 Kriteria Variabel	35
Tabel 4.1 Tipe Data Atribut	39
Tabel 4.2 Hasil Pengujian Data <i>Training</i> dan Data <i>Testing</i>	50
Tabel 4.3 Data <i>Training</i> Penelitian	51
Tabel 4.4 Hasil Perhitungan <i>Information Gain</i> Tertinggi pada <i>Node</i> 1.....	59
Tabel 4.5 Hasil Perhitungan <i>Information Gain</i> Tertinggi pada <i>Node</i> 1.1	66
Tabel 4.6 Hasil Perhitungan <i>Information Gain</i> Tertinggi pada <i>Node</i> 1.1.2.....	73
Tabel 4.7 Hasil Perhitungan <i>Information Gain</i> Tertinggi pada <i>Node</i> 1.1.2.1	79
Tabel 4.8 Hasil Perhitungan <i>Information Gain</i> Tertinggi pada <i>Node</i> 1.1.2.1.1.....	83
Tabel 4.9 Hasil Perhitungan <i>Information Gain</i> Tertinggi pada <i>Node</i> 1.1.2.1.2.....	87
Tabel 4.10 Hasil Perhitungan <i>Information Gain</i> Tertinggi pada <i>Node</i> 1.1.2.1.4...	91
Tabel 4.11 <i>Performance Vector</i>	93

DAFTAR LAMPIRAN

Lampiran I	Lembar Bimbingan Skripsi.....	102
Lampiran II	Surat Ijin Penelitian.....	106
Lampiran III	Formulir dan Hasil Wawancara	107
Lampiran IV	Foto-Foto	110
Lampiran V	Data Training Dan Testing.....	111