

**IMPLEMENTASI ALGORITMA NAÏVE BAYES DALAM PENGKLASIFIKASIAN
STATUS KEPESERTAAN BANTUAN SOSIAL PROGRAM KELUARGA
HARAPAN DI KABUPATEN MANGGARAI TIMUR**

SKRIPSI



disusun oleh

Florentio Freynademetz Saka

17.61.0117

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023**

**IMPLEMENTASI ALGORITMA NAÏVE BAYES DALAM PENGKLASIFIKASIAN
STATUS KEPESERTAAN BANTUAN SOSIAL PROGRAM KELUARGA
HARAPAN DI KABUPATEN MANGGARAI TIMUR**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

Florenthio Freynademetz Saka

17.61.0117

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023**

PERSETUJUAN

SKRIPSI


**IMPLEMENTASI ALGORITMA NAÏVE BAYES DALAM
PENGKLASIFIKASIAN STATUS KEPESERTAAN
BANTUAN SOSIAL PROGRAM KELUARGA
HARAPAN DI KABUPATEN MANGGARAI**

yang dipersiapkan dan disusun oleh

Florenthio Freynademetz Saka
17.61.0117

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 10 Desember 2022

Dosen Pembimbing,


Theopilus Bayu Sasongko, S.Kom, M.Eng.
NIK. 190302375

PENGESAHAN

SKRIPSI

**IMPLEMENTASI ALGORITMA NAÏVE BAYES DALAM
PENGKLASIFIKASIAN STATUS KEPESERTAAN
BANTUAN SOSIAL PROGRAM KELUARGA
HARAPAN DI KABUPATEN MANGGARAI**

yang dipersiapkan dan disusun oleh

Florentio Freynademetz Saka
17.61.0117

telah dipertahankan di depan Dewan Penguji
pada tanggal 21 Desember 2022

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Theopilus Bayu Sasongko, S.Kom, M.Eng.
NIK. 190302375

Uvoek Anggoro Saputro, M.Kom.
NIK. 190302419

Anna Baita, M.Kom.
NIK. 190302290

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, M.Kom.
NIK. 190302096

PERNYATAAN

Saya yang bertanda tangan di bawah ini menyatakan bahwa, skripsi ini adalah hasil karya saya sendiri (ASLI) dan ini skripsi ini belum pernah di terapkan oleh pihak lain untuk menerima gelar akademik di Lembaga Pendidikan manapun, dan sepanjang pengetahuan saya tidak ada karya atau pemikiran yang pernah di tulis dan/atau dipublikasikan oleh siapapun, kecuali yang tertulis yang tercantum dalam naskah ini dan yang mana disebutkan dalam daftar Pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggung jawab saya sendiri.

Yogyakarta,



Florenthio Freynademetz Saka
17.61.0117

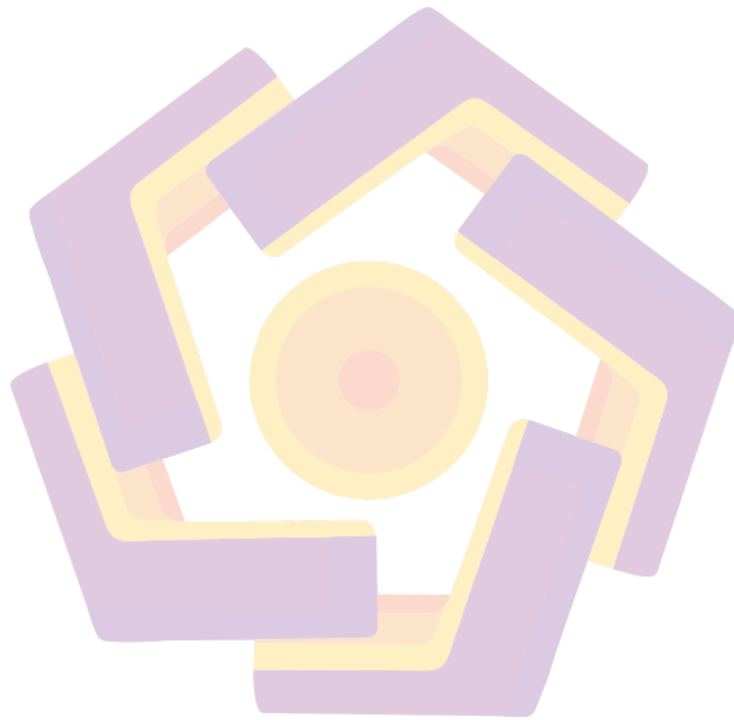
MOTTO

“Kesempatan tidak datang dua kali tapi kesempatan datang kepada siapa yang tidak berhenti mencoba.”

-Dzawin Nur

“It always seems impossible until it's done.”

-Nelson Mandela



PERSEMBAHAN

Dengan rasa syukur yang mendalam, dengan diselesaikannya skripsi ini, skripsi ini saya persembahkan kepada :

1. Kedua orang tua yang selalu ada dan selalu mendukung serta membimbing saya untuk menggapai impian saya. Ibu tersayang, Maria Immaculata Delima yang tidak kenal lelah mengingatkan saya kerasnya dunia sehingga saya harus berusaha sebaik mungkin dengan tidak mengandalkan orang lain, tidak lupa mengajarkan untuk merawat diri, bangga dengan diri sendiri. Untuk Bapakku tersayang, Nikolaus Tolentino Saka, yang tidak kenal lelah untuk
2. Saudara/I terkasih, Gregorio Reynaldo Brechman Saka, Chrysanti Valentine Margaritha, Angelica Prettycia Laveindra, dan Ansgarius Galentino Saka yang selalu memberikan semangat dan dukungan dalam pakerjaan saya.



KATA PENGANTAR

Puji syukur kehadiran Tuhan Yang Maha Esa yang telah memberikan rahmat dan karunia sehingga penulis dapat menyelesaikan skripsi yang berjudul “Implementasi Algoritma Naïve Bayes Dalam Pengklasifikasian Status Kepesertaan Bantuan Sosial Program Keluarga Harapan Di Kabupaten Manggarai”.

Skripsi ini disusun untuk memenuhi syarat perolehan gelar sarjana pada Fakultas Ilmu Komputer, Universitas Amikom, Yogyakarta. Penulis menyadari penyusunan skripsi ini masih jauh dari kata sempurna. Untuk itu penulis sangat mengharapkan saran yang terbaik untuk perbaikan di tahap selanjutnya, juga diharapkan penelitian ini dapat bermanfaat bagi penelitian selanjutnya.

Dengan diselesaikannya skripsi ini, tidak lupa penulis mengucapkan terima kasih kepada :

1. M. Suyanto, Prof, Dr, M.M. selaku Rektor Universitas AMIKOM Yogyakarta
2. Hanif Al Fatta, M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta
3. Erik Hadi Saputra, S.Kom, M.Eng. selaku dosen wali saya
4. Theopilus Bayu Sasongko, S.Kom, M.Eng. selaku dosen pembimbing skripsi yang telah membantu dan memberi pengarahan
5. Segenap dosen pengampu mata kuliah penulis selama berkuliah di Fakultas Ilmu Komputer Universitas Yogyakarta
6. Untuk Bapak, Ibu, Saudara/I yang selalu berdoa ,memberikan dukungan dan semangat
7. Teman – teman BCI 17 yang selalu membantu selama proses belajar
8. BPS Manggarai Timur, Bapeda Manggarai Timur, dan Dinas Sosial Manggarai Timur yang telah membantu selama penulis melakukan penelitian
9. Semua pihak yang telah membantu penulis baik secara langsung dan tidak langsung yang tidak dapat disebutkan satu persatu

Semoga Tuhan membalas kebaikan semua pihak yang telah membantu penulis untuk menyelesaikan penelitian ini.

Yogyakarta,

Florenthio Freynademetz Saka
17.61.0117

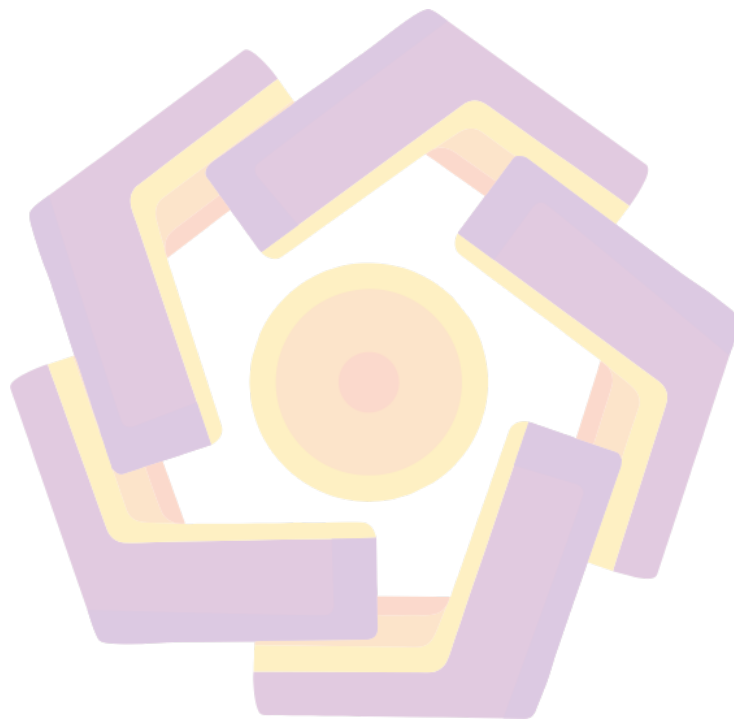
DAFTAR ISI

JUDUL	i
PERSETUJUAN	ii
PENGESAHAN	iii
PERNYATAAN	iv
MOTTO	v
PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
INTISARI	xiii
ABSTRACT	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Maksud dan Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Metodologi Penelitian	4
1.7 Sistematika Penulisan	5
BAB II LANDASAN TEORI	6
2.1 Kajian Pustaka	6
2.2 Landasan Teori	8
2.2.1 Data mining	8
2.2.2 Klasifikasi Data	9
2.2.3 Python	10
2.2.4 Jupyter Notebook	10
2.2.5 Naïve Bayes Classifier	10
2.2.6 Train Test Split	11
2.2.7 Confusion Matrix	12
BAB III METODOLOGI	14
3.1 Alur Penelitian	14

3.2 Model Proses CRISP-DM	14
3.2.1 Business Understanding	14
3.2.2 Data Understanding	15
3.2.3 Data Preparation	17
3.2.4 Tahap Modeling	23
3.2.5 Evaluasi	23
BAB IV HASIL DAN PEMBAHASAN	24
4.1 Import Dataset	24
4.2 Preprocessing Data	25
4.2.1 Pembersihan Data	25
4.3 Label Encoder	30
4.4 Rescalling	31
4.5 Pembagian Data Latih Dan Uji	31
4.6 Implementasi Algoritma Naïve Bayes	32
4.7 Pengujian Dengan Confusion Matrix	33
4.8 Resampling	34
4.9 Implementasi	36
BAB V PENUTUP	39
5.1 Kesimpulan	39
5.2 Saran	39
DAFTAR PUSTAKA	40

DAFTAR TABEL

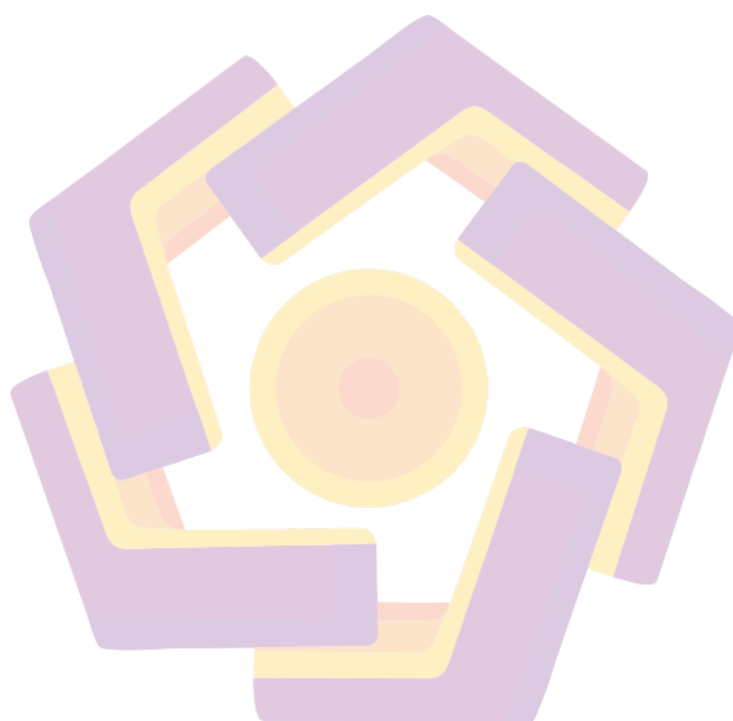
Tabel 2.1 Kajian Pustaka	7
Tabel 2.2 Confusion Matrix.....	13
Tabel 3.1 Atribut Data	15
Tabel 3.2 Atribut Data Yang Dipilih	19
Tabel 4.1 Confusion Matix	33



DAFTAR GAMBAR

Gambar 2. 1 Train Test Split	12
Gambar 3.1 Alur Penelitian	14
Gambar 3.2 Missing Value	20
Gambar 3.3 Duplicate Data	21
Gambar 3.4 Data Ketegorical	22
Gambar 4.1 File BDT	24
Gambar 4.2 Import Dataset.....	24
Gambar 4.3 Hasil Import Dataset	25
Gambar 4.4 Jumlah Missing Value.....	26
Gambar 4.5 Fungsi Menghapus Data Null	26
Gambar 4.6 Hasil Menghapus Missing Value	27
Gambar 4.7 Informasi Dataset Sebelum Drop Null.....	27
Gambar 4.8 Informasi Dataset Setelah Drop Null.....	28
Gambar 4.9 Duplikasi Data	28
Gambar 4.10 Fungsi Menghilnagkan Duplikasi Data	29
Gambar 4.11 Dataset Tanpa Duplikasi Data	29
Gambar 4.12 Informasi Dataset	30
Gambar 4.13 Fungsi Label Encoder	30
Gambar 4.14 Label Sebelum Dikonversi.....	31
Gambar 4.15 Label Setelah Dikonversi	31
Gambar 4. 16 Min Max Scaller	31
Gambar 4.17 Train Test Split	31
Gambar 4.18 X_train	32
Gambar 4.19 X_test	32
Gambar 4.20 y_train	32
Gambar 4.21 y_test	32
Gambar 4.22 Fungsi Naïve Bayes Classifier	33
Gambar 4.23 Visualisasi Confusion Matix.....	34
Gambar 4.24 <i>Classification Report</i>	34
Gambar 4. 25 Unbalanced Data	35
Gambar 4. 26 Fungsi Under Sampling	35
Gambar 4. 27 Pembagian Data Setelah Diresampling.....	35

Gambar 4. 28 Classification report..... 36
Gambar 4.29 Form Data 36
Gambar 4.30 Hasil klasifikasi Data 38



INTISARI

Program Keluarga Harapan yang selanjutnya disebut PKH adalah program pemberian bantuan sosial bersyarat kepada Keluarga Miskin (KM) yang ditetapkan sebagai keluarga penerima manfaat PKH. Sebagai upaya percepatan penanggulangan kemiskinan, sejak tahun 2007 Pemerintah Indonesia telah melaksanakan PKH. Program Perlindungan Sosial yang juga dikenal di dunia internasional dengan istilah *Conditional Cash Transfers* (CCT) telah terbukti cukup berhasil dalam menanggulangi kemiskinan yang dihadapi oleh negara-negara miskin dan berkembang. Sebagai sebuah program bantuan sosial bersyarat, PKH membuka akses keluarga miskin terutama ibu hamil dan anak untuk memanfaatkan berbagai fasilitas layanan kesehatan (faskes) dan fasilitas layanan pendidikan (fasdik) yang tersedia di sekitar mereka. Manfaat PKH juga mulai didorong untuk mencakup penyandang disabilitas dan lanjut usia dengan mempertahankan taraf kesejahteraan sosialnya sesuai dengan amanat konstitusi.

Sebagai suatu program bantuan sosial bersyarat dari pemerintah untuk membantu masyarakat yang dianggap kurang mampu, sering dijumpai penetapan sasaran penerima PKH belum dilakukan dengan baik. Dalam penerapannya PKH masih belum maksimal dikarenakan kurang tepat sasaran bagi penerima bantuan sosial ini. Oleh karena itu, penelitian ini bertujuan untuk mengklasifikasi status keikutsertaan dalam bansos PKH menggunakan algoritma *Naïve bayes*.

Data mining dapat membantu untuk mempermudah mencari solusi masalah ini. Hasil implementasi menggunakan algoritma *Naïve Bayes* serta metode *CRIPS-DM* dengan menggunakan data penduduk dari Kabupaten Manggarai Timur adalah sebesar 60%.

Kata Kunci : Naïve Bayes, PKH, Klasifikasi, Data Mining, CRIPS-DM

ABSTRACT

Program Keluarga Harapan, hereinafter referred to as PKH, is a conditional social assistance program for Poor Families who are designated as PKH beneficiary families. As an effort to accelerate poverty alleviation, since 2007 the Government of Indonesia has implemented PKH. The Social Protection Program, also known internationally as Conditional Cash Transfers (CCT), has proven to be quite successful in overcoming poverty faced by poor and developing countries. As a conditional social assistance program, PKH opens access for poor families, especially pregnant women and children, to take advantage of the various health service facilities and education service facilities available around them. PKH benefits have also begun to be encouraged to include persons with disabilities and the elderly by maintaining their level of social welfare in accordance with the constitutional mandate.

As a conditional social assistance program from the government to help people who are considered less fortunate, it is often found that targeting PKH recipients has not been carried out properly. In its implementation, PKH is still not optimal because it is not well targeted for recipients of this social assistance. Therefore, this study aims to classify participation status in PKH social assistance using the Naïve Bayes algorithm.

Data mining can help to make it easier to find solutions to this problem. The results of the implementation using the Naive Bayes algorithm and the CRIPS-DM method using population data from East Manggarai Regency were 60%.

Keywords: Naïve Bayes, PKH, Classification, Data Mining, CRIPS-DM

