

**ANALISIS TINGKAT AKURASI ALGORITMA K-NEAREST
NEIGHBOUR DENGAN NAÏVE BAYES UNTUK
KLASIFIKASI KELAYAKAN
PENDONOR DARAH
(Studi Kasus : RSUD Unit Transfusi Darah Bima)**

SKRIPSI



disusun oleh

SITI MUTMAINAH

18.21.1146

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2020**

**ANALISIS TINGKAT AKURASI ALGORITMA K-NEAREST
NEIGHBOUR DENGAN NAÏVE BAYES UNTUK
KLASIFIKASI KELAYAKAN
PENDONOR DARAH
(Studi Kasus : RSUD Unit Transfusi Darah Bima)**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Infotmatika



disusun oleh

SITI MUTMAINAH

18.21.1146

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2020**

PERSETUJUAN

SKRIPSI

**ANALISIS TINGKAT AKURASI ALGORITMA K-NEAREST
NEIGHBOUR DENGAN NAÏVE BAYES UNTUK
KLASIFIKASI KELAYAKAN
PENDONOR DARAH
(Studi Kasus : RSUD Unit Transfusi Darah Bima)**

yang dipersiapkan dan disusun oleh

SITI MUTMAINAH

18.21.1146

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 14 Mei 2019

Dosen Pembimbing,



Mardhiva Hayatv, S.T., M.Kom

NIK. 190302108

PENGESAHAN

SKRIPSI

**ANALISIS TINGKAT AKURASI ALGORITMA K-NEAREST
NEIGHBOUR DENGAN NAÏVE BAYES UNTUK**

KLASIFIKASI KELAYAKAN

PENDONOR DARAH

(Studi Kasus : RSUD Unit Transfusi Darah Bima)

yang dipersiapkan dan disusun oleh

SITI MUTMAINAH

18.21.1146

telah dipertahankan di depan Dewan Penguji
pada tanggal 10 Febuari 2020

Susunan Dewan Penguji

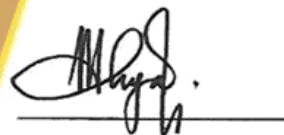
Nama Penguji

Hartatik, S.T., M.Cs
NIK. 190302232

Yuli Astuti, M.Kom
NIK. 190302146

Mardhiya Hayaty, S.T., M.Kom
NIK. 190302108

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal Febuari 2020

DEKAN FAKULTAS ILMU KOMPUTER

Krisnawati, S.Si, M.T.
NIK. 190302038

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 12-02-20



SITI MUTMAINAH
NIM. 18.21.1146

MOTTO

QS. AL-ANKABUT: AYAT 6 :

يُجَاهِدُ فَايَّمَا جَاهِدَ وَمَنْ لِنَفْسِهِ (6)

"Barang siapa yang bersungguh sungguh, sesungguhnya kesungguhan tersebut untuk kebaikan dirinya sendiri"

QS. Asy Syarh ayat 5-6,

اَيْسَرَ لُغُسْرًا مَعَ فَايَنَّ (5) يُسْرًا لُغُسْرًا مَعَ اِيَنَّ (6)

"Maka sesungguhnya sesudah kesulitan itu ada kemudahan, Sesungguhnya sesudah kesulitan itu ada kemudahan"

PERSEMBAHAN

Puji syukurku kepada Allah SWT yang Maha Esa yang selalu melimpahkan rahmat dan kesehatan sehingga dapat terselesainya skripsi ini. Setelah terabaikan selama satu semester, skripsi ini pun terselesaikan, Skripsi ini akan saya persembahkan kepada :

1. Orang tuaku tercinta, terima kasih yang sebesar-besarnya telah memberikan seluruh kasih sayangnya, serta senantiasa mendoakan agar selalu dimudahkan dalam semua urusanku. Terima kasih banyak mama, papa.
2. Mardhiya Hayaty, S.T., M.Kom selaku dosen pembimbing skripsi ini, terima kasih bu telah membimbing dan menyemangati. Skripsi ini tidak akan selesai tanpa bantuan ibuk.
3. Teman-temanku seperjuangan : Rini, Pai, Yoga, Adit, Dwi, Adi biar sweet sebut nama. Makasih telah sahabat dan teman yang baik selama setahun lebih ini, walaupun gak baik banget sih.
4. Go-Jek dan Grab, Terima kasih telah mengantar kuliah dan makanan pada saat mengerjakan tugas dan skripsi ini disaat hujan, badai, macet dan panas. Kalian adalah pahlawan disaat perut kosong tengah malam.

KATA PENGANTAR

Puji syukur kepada tuhan yang maha esa berkat rahmat dan hidayahnya, hingga penulis dapat menyelesaikan Skripsi dengan judul : “Analisis Tingkat Akurasi Algoritma K-Nearest Neighbour dengan Naïve Bayes untuk Klasifikasi Kelayakan Pendor Donor Darah (Studi Kasus : RSUD Unit Transfusi Darah Bima)” . Skripsi ini bertujuan untuk memenuhi salah satu syarat guna memperoleh gelar Sarjana Komputer (S.Kom) jurusan S1 Informatika. Dalam penulisan Skripsi ini, banyak pihak yang terlibat dan membantu, membina dan membimbing dalam proses menyelesaikan Skripsi ini. penulis menyampaikan banyak-banyak terima kasih kepada :

1. Bapak Prof. Dr. M.Suyanto, MM selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Krisnawati, S.Si, M.T selaku ketua program studi S1 informatika.
3. Ibu Mardhiya hayaty, S.T., M.Kom selaku dosen pembimbing dalam Skripsi.
4. Orang tua yang selalu mendoakan dan mendorong penulis sehingga Skripsi ini dapat diselesaikan dengan baik.
5. Seluruh Bapak dan Ibu dosen, khususnya yang pernah mengajar pada kelas 18 Informatika Transfer.

Yogyakarta, 26 Februari 2020

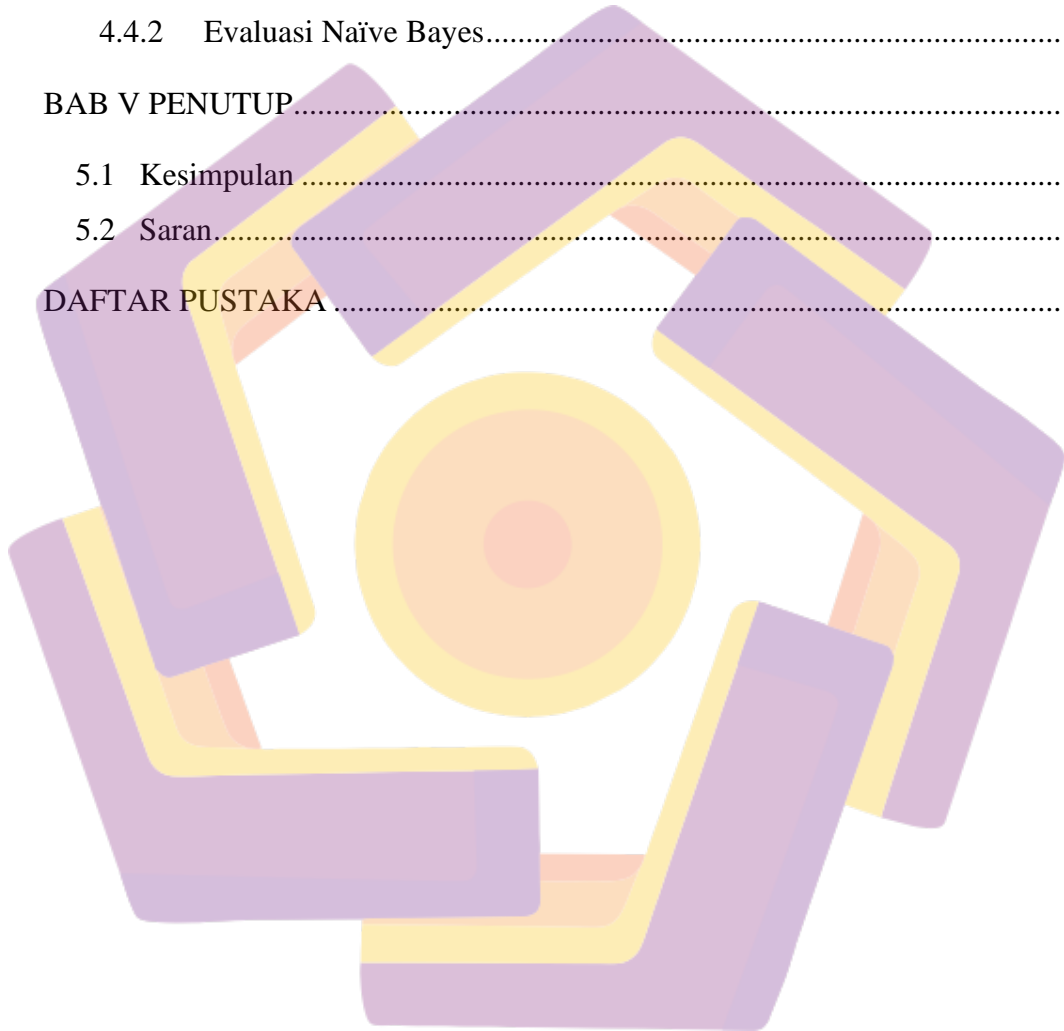
Siti Mutmainah

DAFTAR ISI

COVER	I
JUDUL	II
PERSETUJUAN	III
PENGESAHAN	IV
PERNYATAAN	V
MOTTO	VI
PERSEMBAHAN	VII
KATA PENGANTAR	VIII
DAFTAR ISI	IX
DAFTAR TABEL	XII
DAFTAR GAMBAR	XIII
INTISARI	XV
ABSTRACT	XVI
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Maksud dan Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Metode Penelitian	4
1.7 Sistematika Penulisan	4
BAB II LANDASAN TEORI	6
2.1 Tinjauan Pustaka	6

2.2	Dasar Teori.....	10
2.2.1	Data Mining	10
2.2.1.1	Proses Data Mining	10
2.2.2	Dataset.....	11
2.2.3	Klasifikasi	15
2.2.4	Naïve Bayes	17
2.2.5	K-Nearest Neighbour	18
2.2.6	SMOTE (Sythetic Minority Over-sampling Techniqu).....	22
2.2.7	Pengukuran Kinerja Klasifikator (Matriks Confusion).....	23
2.2.7.1	Akurasi	24
2.2.7.2	Sensitifitas dan Spesifisitas	25
2.2.7.3	Precision dan Recall.....	26
BAB III METODE PENELITIAN.....		27
3.1	Alat dan Bahan Penelitian.....	27
3.1.1	Alat Penelitian.....	27
3.1.2	Data Penelitian	27
3.2	Langkah Penelitian.....	29
3.2.1	Pengumpulan Dataset.....	30
3.2.2	Preprocessing/Pengolahan Data	31
3.2.3	Pemodelan.....	31
3.2.4	Evaluasi.....	32
BAB IV HASIL DAN PEMBAHASAN		33
4.1	Dataset.....	33
4.2	Praprocessing / Pengolahan data.....	36
4.2.1	Cek Missing Data.....	36
4.2.2	Penambahan data.....	37
4.2.3	Transformasi data.....	38
4.2.4	Penanganan Imbalance class	42

4.3	Pemodelan.....	43
4.3.1	Klasifikasi K-Nearest Neigbord.....	43
4.3.2	Klasifikasi Naïve Bayes	43
4.4	Evaluasi.....	47
4.4.1	Evaluasi K-Nearest Neigbord	47
4.4.2	Evaluasi Naïve Bayes.....	48
BAB V PENUTUP.....		51
5.1	Kesimpulan	51
5.2	Saran.....	52
DAFTAR PUSTAKA		53



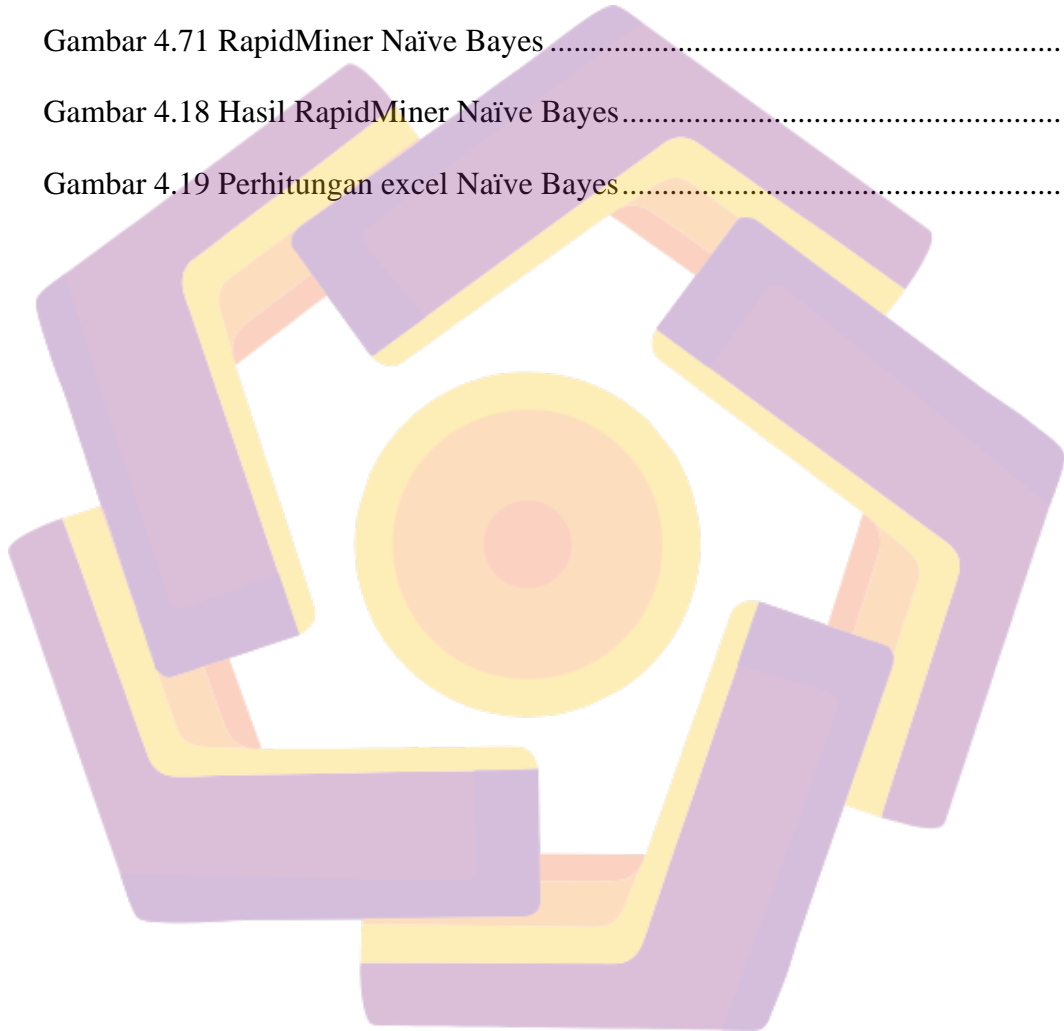
DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka	8
Tabel 2.2 Lanjutan Tinjauan Pustaka.....	9
Tabel 2.3 Matriks confusion	24
Tabel 2.4 Matriks confusion Sensitifitas.....	26
Tabel 4.1 Dataset.....	33
Tabel 4.2 Lanjutan tabel.....	33
Tabel 4.3 Atribut	36
Tabel 4.4 Setelah penanganan Missing Value.....	38
Tabel 4.5 Data pendonor layak dan tidak.....	41
Tabel 4.6 Proporsi data Train dan Testing	41
Tabel 4.7 Proporsi data SMOTE.....	42
Tabel 4.8 Matrik Confision KNN dan SMOTE	43
Tabel 4.9 K=1 sampai k=10.....	44
Tabel 4.10 Matrik Confision NB dan SMOTE	46
Tabel 4.11 Matrik Confision KNN	47
Tabel 4.12 Evaluasi Matrik Confision SMOTE KNN	49
Tabel 4.13 Evaluasi Matrik Confision NB	50
Tabel 4.14 Matrik Confision NB DAN SMOTE	52
Tabel 4.15 PERBANDINGAN	52

DAFTAR GAMBAR

Gambar 2.1 Proses Kerja Klasifikasi	16
Gambar 2.2 K-NN Tetangga terdekat 1	19
Gambar 2.3 K-NN Tetangga terdekat 3	19
Gambar 2.4 K-NN Tetangga terdekat 5	20
Gambar 2.5 K-NN Tetangga terdekat 7	20
Gambar 2.6 Alur K-NN.....	21
Gambar 2.7 Proses SMOTE.....	23
Gambar 3.1 Alur Donor Darah.....	29
Gambar 3.2 Langkah Penelitian.....	30
Gambar 3.3 Pemodelan	32
Gambar 4.1 Dataset.....	35
Gambar 4.2 Source Code Sistole	37
Gambar 4.3 Source Code Rata-rata.....	37
Gambar 4.4 Source Code Nadi.....	38
Gambar 4.5 Dataset Numerik.....	40
Gambar 4.6 Data berdasarkan Class	41
Gambar 4.7 Source Code SMOTE.....	42
Gambar 4.8 KNN Non SMOTE.....	43
Gambar 4.9 KNN SMOTE.....	43
Gambar 4.10 Grafik akurasi KNN	45
Gambar 4.11 Grafik akurasi SMOTE	45
Gambar 4.12 Naïve bayes	46

Gambar 4.13 Naïve Bayes dan SMOTE	46
Gambar 4.14 RapidMiner KNN	48
Gambar 4.15 Hasil RapidMiner KNN	48
Gambar 4.16 Perhitungan excel KNN	48
Gambar 4.71 RapidMiner Naïve Bayes	51
Gambar 4.18 Hasil RapidMiner Naïve Bayes.....	51
Gambar 4.19 Perhitungan excel Naïve Bayes.....	51



INTISARI

RSUD atau Rumah Sakit Umum Daerah Unit Transfusi Darah menyelenggarakan donor darah.. sebelum melakukan donor darah pendonor harus melakukan pemeriksaan. Klasifikasi merupakan salah satu teknik data mining yang dapat melakukan prediksi data pendonor dalam menentukan kelayakan pendonor darah.

Penelitian ini bermaksud untuk membandingkan akurasi antara Algoritma K-Nearest Neigbord dan Naïve Bayes dalam klasifikasi kelayakan pendonor, dataset yang didapat pada rumah sakit umum daerah Bima memiliki *imbalance class* antara pendonor (Layak) dan (Tidak), sehingga data yang ada harus dilakukan penanganan imbalance data. Teknik yang digunakan dalam menangani *Imbalance class* adalah SMOTE (*Syntetic Minority Over Sampling Technique*).

Penelitian dengan K-Nearest Neigbord menerapkan nilai K yang berbeda yaitu 1 sampai 10 untuk melihat hasil kinerja yang paling tinggi, akurasi tertinggi K-Nearest Neigbord pada K=1 yaitu 100%, mengalami penurunan akurasi pada K=2,3 92.13% dan K=4 sampai 10 stabil pada 91.17%. Penerapan SMOTE dalam menangani imbalance class pada klasifikasi pendonor darah menghasilkan akurasi tertinggi pada K=2 yaitu 98.09%. Naïve Bayes memiliki akurasi 98.09%. sedangkan akurasi yang sama antara Naïve Bayes dan Naïve Bayes dengan SMOTE 99.04%.

Kata Kunci: Klasifikasi, Pendonor darah, K-Nearest Neigbord, Naïve Bayes, SMOTE, Confusion matrix.

ABSTRACT

Hospital or Regional Blood Transfusion Unit Hospital organizes blood donation to maintain blood supply, so that when blood is needed it is available. But not all donors can donate their blood, before donating blood donors have to do an examination. Classification is a data mining technique that can predict donor data in determining the eligibility of blood donors.

This study intends to compare the accuracy between the K-Nearest Neighbord Algorithm and Naïve Bayes in the donor eligibility classification, the dataset obtained at the Bima area general hospital has a imbalance class between donors (Eligible) and (No), so that the existing data must be done handling imbalance data. The technique used in handling Imbalance class is SMOTE (Syntetic Minority Over Sampling Technique).

Research with K-Nearest Neighbord applies different K values, namely 1 to 10 to see the highest performance results, the highest accuracy of K-Nearest Neighbord at $K = 1$ is 100%, has decreased accuracy at $K = 2$ 92.13% and $K = 4$ to 10 stable at 91.17%. The application of SMOTE in handling imbalance class in the classification of blood donors produced the highest accuracy at $K = 2$, namely 98.09%. Naïve Bayes has an accuracy of 98.09%. while the same accuracy between Naïve Bayes and Naïve Bayes + SMOTE 99.04%.

Keyword: *Classification, Blood donor, K-Nearest Neighbord, Naïve Bayes, SMOTE, Confusion matrix.*