

TUGAS AKHIR

**PERBANDINGAN ALGORITMA *NAIVE BAYES* DENGAN
ALGORITMA *K-NEAREST NEIGHBOR* UNTUK
KLASIFIKASI PENYAKIT JANTUNG**



Oleh :

AGIL LANGGA DIKAYANTO

1410651223

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER**

2019

TUGAS AKHIR

**PERBANDINGAN ALGORITMA *NAIVE BAYES* DENGAN
ALGORITMA *K-NEAREST NEIGHBOR* UNTUK
KLASIFIKASI PENYAKIT JANTUNG**

Diajukan Untuk Memenuhi Persyaratan Guna Meraih Gelar Sarjana Komputer
Teknik Informatika Universitas Muhammadiyah Jember



Oleh :

AGIL LANGGA DIKAYANTO

1410651223

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER**

2019

HALAMAN PENGESAHAN

**PERBANDINGAN ALGORITMA *NAIVE BAYES* DENGAN
ALGORITMA *K-NEAREST NEIGHBOR* UNTUK KLASIFIKASI
PENYAKIT JANTUNG**

Oleh:

AGIL LANGGA DIKAYANTO

14 1065 1223

Telah mempertanggung jawabkan Laporan Tugas Akhirnya pada sidang Tugas Akhir tanggal 26 Maret 2019 sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom)

di

Universitas Muhammadiyah Jember

Disetujui Oleh :

Pembimbing

Agung Nilogiri S.T., M.Kom.

NIP. 19770330 200501 1 002

Penguji I

Penguji II

Deni Arifianto, M.Kom.

NPK. 11 03 588

Mengesahkan,

Dekan Fakultas Teknik

Triawan Adi Cahyanto, M.Kom.

NPK. 12 03 719

Mengetahui,

Ketua Program Studi Teknik Informatika

Ir. Suhartinah, MT

NPK. 95 05 246

Yeni Dwi Rahayu, S.ST, M.Kom.

NPK. 11 03 590

KATA PENGANTAR

Alhamdulillah kami panjatkan rasa syukur kepada Allah SWT .karena hanya dengan rahmat, hidayah dan inayah-Nya kami dapat menyelesaikan Tugas Akhir ini dengan judul :

PERBANDINGAN ALGORITMA NAIVE BAYES DENGAN ALGORITMA K-NEAREST NEIGHBOR UNTUK KLASIFIKASI PENYAKIT JANTUNG

Dalam menyelesaikan Tugas Akhir ini, kami berpegang pada teori yang pernah kami dapatkan dan bimbingan dari dosen pembimbing Tugas Akhir dan pihak-pihak lain yang sangat membantu hingga sampai terselesaikannya Tugas Akhir ini. Tugas Akhir ini merupakan salah satu syarat akademis untuk memperoleh gelar Sarjana Komputer (S.Kom) di Fakultas Teknik, Universitas Muhammadiyah Jember.

Kami menyadari bahwa masih banyak kekurangan pada analisis dan pembuatan buku Tugas Akhir ini. Oleh karena itu, besar harapan kami untuk menerima saran dan kritik dari para pembaca. Semoga buku ini dapat memberikan manfaat bagi para mahasiswa Universitas Muhammadiyah Jember pada umumnya dan dapat memberikan nilai lebih untuk para pembaca pada khususnya.

Jember, 26 Maret 2019

Penulis

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR	iii
HALAMAN PENGESAHAN	iii
PERNYATAAN KEASLIAN	iv
ABSTRAK	v
ABSTRACT	vi
HALAMAN PERSEMBAHAN	vii
UNGKAPAN TERIMA KASIH	viii
KATA PENGANTAR	x
MOTTO	xi
DAFTAR ISI	xii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvi
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan Penelitian.....	3
1.4. Batasan Masalah.....	3
1.5. Manfaat Penelitian.....	3
BAB II TINJAUAN PUSTAKA	5
2.1. Penelitian Terdahulu	5
2.2. Data Mining.....	6
2.2.1. Tahapan- Tahapan Data Mining	6
2.3. Klasifikasi.....	8
2.3.1. Konsep Klasifikasi	8
2.3.2. Model	8
2.3.3. Pengukuran Klasifikasi	10
2.3.4. K-Fold Cross Validation	12

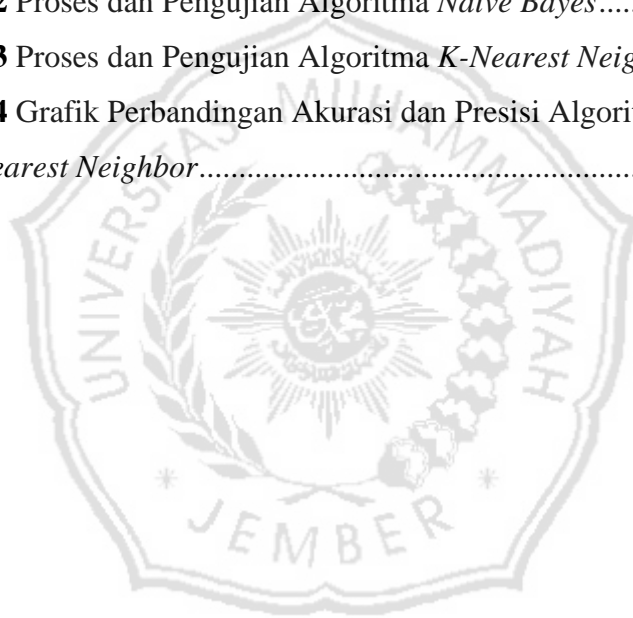
2.4. Pengertian Analisis.....	13
2.5. Naive Bayes.....	14
2.5.1. Teorema Bayes.....	14
2.5.2. <i>Naive Bayes</i> untuk Klasifikasi	17
2.6. K-Nearest Neighbor	18
2.6.1. Karakteristik Klasifikasi dengan <i>K-Nearest Neighbor</i>	21
2.7. Penyakit Jantung	21
2.7.1. Faktor Resiko Penyakit Jantung.....	22
2.7.2. Gejala Penyakit Jantung.....	27
2.8. RapidMiner Studio	28
BAB III METODE PENELITIAN	31
3.1. Studi Literatur	31
3.2. Pengumpulan Data	32
3.3. Preprocessing Data.....	33
3.4. Model	33
3.5. Implementasi	34
3.5.1. Perhitungan Algoritma <i>Naive Bayes</i>	35
3.5.2. Perhitungan Algoritma <i>K-Nearest Neighbor</i>	38
3.5. Pengujian.....	41
3.6. Analisis.....	41
3.7. Kesimpulan.....	42
BAB IV HASIL DAN PEMBAHASAN	43
4.1. Implementai dengan RapidMiner Studio	43
4.1.1. Implementasi Algoritma <i>Naive Bayes</i>	45
4.1.2. Implementasi Algoritma <i>K-Nearest Neighbor</i>	46
4.2. Pengujian Algoritma <i>Naive Bayes</i>	48
4.3. Pengujian Algoritma <i>K-Nearest Naighbor</i>	50
4.4. Analisis.....	57
BAB V KESIMPULAN DAN SARAN	61
1.1. Kesimpulan.....	61
1.2. Saran.....	61

DAFTAR PUSTAKA	63
LAMPIRAN 1.....	66
LAMPIRAN 2.....	80
LAMPIRAN 3.....	82



DAFTAR GAMBAR

Gambar 2. 1 Proses Data Mining	7
Gambar 2. 2 Proses Pekerjaan Klasifikasi.....	9
Gambar 2. 3 KNN dengan nilai K-tetangga (a) 1-NN, (b) 2-NN, (c) 3-NN, (d) 7-NN.....	20
Gambar 3. 1 Tahap Penelitian	31
Gambar 3. 2 Algoritma Metode (a) <i>Naive Bayes</i> , (b) <i>K-Nearest Neighbor</i>	34
Gambar 4. 1 Tampilan Teknik Pengujian <i>Cross Validation</i>	43
Gambar 4. 2 Proses dan Pengujian Algoritma <i>Naive Bayes</i>	45
Gambar 4. 3 Proses dan Pengujian Algoritma <i>K-Nearest Neighbor</i>	46
Gambar 4. 4 Grafik Perbandingan Akurasi dan Presisi Algoritma <i>Naive Bayes</i> dengan <i>K-Nearest Neighbor</i>	60



DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu.....	5
Tabel 2. 2 Matriks Konfusi untuk Klasifikasi Dua Kelas	11
Tabel 2. 3 Contoh Dataset	13
Tabel 2. 4 Contoh Data Eksperimen	13
Tabel 2. 5 Penjelasan Formula <i>Naive Bayes</i>	15
Tabel 2. 6 Konversi Kolesterol.....	23
Tabel 2. 7 Konversi Tekanan Darah.....	23
Tabel 2. 8 Konversi Tekanan Jantung.....	24
Tabel 2. 9 Konversi Kadar Gula Darah	24
Tabel 2. 10 Konversi Usia.....	25
Tabel 2. 11 Konversi Jenis Kelamin.....	25
Tabel 2. 12 Koneversi Elektrokardiografi	26
Tabel 2. 13 Konversi Oldpeak.....	26
Tabel 2. 14 Konversi Slope	27
Tabel 2. 15 Konversi Scan Thallium.....	27
Tabel 2. 16 Konversi Jenis Sakit Dada.....	27
Tabel 2. 17 Konversi Angina Induksi.....	28
Tabel 3. 1 Atribut dan Data Penyakit Jantung.....	32
Tabel 3. 2 Missing Data	33
Tabel 3. 3 Data Training <i>Naive Bayes</i>	35
Tabel 3. 4 Data Testing <i>Naive Bayes</i>	35
Tabel 3. 5 Jumlah Kelas Data Training <i>Naive Bayes</i>	35
Tabel 3. 6 Hitung Jumlah Kelas yang Sama dengan Kelas yang Sama	36
Tabel 3. 7 Data Testing <i>Naive Bayes</i>	37
Tabel 3. 8 Klasifikasi <i>Naive Bayes</i>	38
Tabel 3. 9 Data Training <i>K-Nearest Neighbor</i>	39
Tabel 3. 10 Data Testing <i>K-Nearest Neighbor</i>	39
Tabel 3. 11 Hasil Perhitungan <i>Euclidian Distance</i>	40
Tabel 3. 12 Ranking Jarak <i>Euclidian Distance</i>	40

Tabel 3. 13 Klasifikasi Algoritma <i>K-Nearest Neighbor</i>	41
Tabel 3. 14 Rencana Perbandingan <i>Naive Bayes & K-Nearest Neighbor</i>	41
Tabel 4. 1 Hasil <i>Confussion Matrix</i> Algoritma <i>Naive Bayes</i> dengan <i>2-fold Cross Validation</i> pada Folder Data Uji ke-1	46
Tabel 4. 2 Hasil <i>Confussion Matrix</i> Algoritma <i>K-Nearest Neighbor</i> dengan 3 Ketetanggaan Menggunakan <i>2-fold Cross Validation</i> pada Folder Data Uji ke-1 48	
Tabel 4. 3 Hasil Pengujian Algoritma <i>Naive Bayes</i>	48
Tabel 4. 4 Hasil Pengujian Algoritma <i>K-Nearest Neighbor</i> dengan 3 Ketetanggaan.....	50
Tabel 4. 5 Hasil Pengujian Algoritma <i>K-Nearest Neighbor</i> dengan 5 Ketetanggaan.....	52
Tabel 4. 6 Pengujian Algoritma <i>K-Nearest Neighbor</i> dengan 7 Ketetanggaan ...	54
Tabel 4. 7 Pengujian Algoritma <i>K-Nearest Neighbor</i> dengan 9 Ketetanggaan ...	55
Tabel 4. 8 Akurasi Tertinggi <i>K-Nearest Neighbor</i> dengan 3 Ketetanggaan	57
Tabel 4. 9 Akurasi Tertinggi <i>K-Nearest Neighbor</i> dengan 5 Ketetanggaan	58
Tabel 4. 10 Akurasi Tertinggi <i>K-Nearest Neighbor</i> dengan 7 Ketetanggaan	58
Tabel 4. 11 Akurasi Tertinggi <i>K-Nearest Neighbor</i> dengan 9 Ketetanggaan	58
Tabel 4. 12 Hasil Akurasi Tertinggi <i>K-Nearest Neighbor</i> dari Masing-masing Jarak Ketetanggaan	59
Tabel 4. 13 Hasil Perbandingan Algoritma <i>Naive Bayes</i> dengan <i>K-Nearest Neighbor</i>	59

DAFTAR PUSTAKA

- Aprilla C, D., Aji Baskoro, D., Ambarwati, L., & Wayan Simri Wicaksana, I. (2013). *Belajar Data Mining dengan RapidMiner*. (R. Sanjaya, Ed.). Jakarta: Open Content Model.
- Arisandy, D. L. (2017). *Analisis Perbandingan Algoritma Naive Bayes dan Algoritma C4.5 Untuk Klasifikasi Multi Data*.
- Health, H. (2018). Reading The New Blood Pressure Guidelines. Retrieved March 28, 2019, from <https://www.health.harvard.edu/heart-health/reading-the-new-blood-pressure-guidelines>
- Hospital Authority. (2016). Coronary Heart Disease Indonesian. *Smart Pasien*.
- Huda, N. M. (2010). *Aplikasi Data Mining Untuk Menampilkan Informasi Tingkat Kelulusan Mahasiswa (Studi Kasus di Fakultas MIPA Universitas Diponegoro)*. Universitas Diponegoro.
- Kusrini, & Luthfi, E. T. (2009). *Algoritma Data Mining*. Yogyakarta: Andi Offset.
- Medistra. (2017). Apakah Penyebab Dari Penyakit Jantung Koroner? Retrieved April 11, 2018, from https://www.medistra.com/index.php?option=com_content&view=article&id=76
- Muamalatku. (2018). Kategori Umur Menurut WHO yang Belum Banyak Diketahui Masyarakat. Retrieved March 28, 2019, from <https://muamalatku.com/kategori-umur-menurut-who/>
- Mustakim, & Oktaviani, G. (2016). Algoritma K-Nearest Neighbor Classification Sebagai Sistem Prediksi Predikat Prestasi Mahasiswa. *Jurnal Sains, Teknologi Dan Industri*, 13(2).
- Novianti, A. G., & Prasetyo, D. (2017). Penerapan Algoritma K-Nearest Neighbor (K-NN) untuk Prediksi Waktu Kelulusan Mahasiswa. *Seminar Nasional APTIKOM*.
- Nurhayati, H., & Nugroho, F. (2012). Implementasi Fuzzy Expert System Untuk Diagnosis Penyakit Jantung. *Implementasi Fuzzy Expert System Untuk Diagnosis Penyakit Jantung*.
- Oemiati, R., & Rustika. (2015). Faktor Risiko Penyakit Jantung Koroner (PJK)

- Pada Perempuan (Baseline Studi Kohor Faktor Risiko PTM). *Buletin Penelitian Sistem Kesehatan*, 18(1), 47–55.
- Oktaviana, A. R. (2016). *Penerapan Data Mining Klasifikasi Pola Nasabah Menggunakan Algoritma C4.5 Pada Bank BRI Batang*. Universitas Dian Nuswantoro.
- Prasetyo, E. (2012). *Data Mining Konsep dan Aplikasi Menggunakan MATLAB*. Yogyakarta: Andi Offset.
- Rifai, B. (2013). Algoritma Neural Network untuk Prediksi Penyakit Jantung. *Techno Nusa Mandiri*, IX(1), 1–9.
- Rohman, A. (2012). Model Algoritma K-Nearest Neighbor (K-NN) untuk Prediksi Kelulusan Mahasiswa. *Dosen Jurusan Elektronika Fakultas Teknik Universitas Pandanaran Semarang*.
- Samiadi, L. A. (2018). Penyakit Angina. Retrieved December 8, 2018, from <https://helohehat.com/penyakit/angina/>
- Sani, R. R., Zeniarja, J., & Luthfiarta, A. (2016). Penerapan Algoritma K-Nearest Neighbor pada Information Retrieval dalam Penentuan Topik Referensi Tugas Akhir. *Journal of Applied Intelligent System*, 1(2).
- Sartika, D., & Sensuse, D. I. (2017). Perbandingan Algoritma Klasifikasi Naive Bayes , Nearest Neighbour , dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian. *Jatisi*, 1(2).
- Sehat, D. (2017). Jantung Koroner – Penyebab, Gejala, Diagnosa, Pengobatan, dan Pencegahan. Retrieved April 11, 2018, from <http://doktersehat.com/jantung-koroner/>
- Solichin, A. (2017). Mengukur Kinerja Algoritma Klasifikasi dengan Confusion Matrix. Retrieved December 8, 2018, from <http://achmatim.net/2017/03/19/mengukur-kinerja-algoritma-klasifikasi-dengan-confusion-matrix/>
- Sumarlin. (2015). Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung Keputusan Klasifikasi Penerima Beasiswa PPA dan BBM. *Jurnal Sistem Informasi Bisnis*, 01.
- Trilipix. (2019). What are the cholesterol numbers ? Retrieved March 28, 2019,

from <https://trilipix.com/what-are-the-cholesterol-numbers>

Wardhani, R. S. (2014). Aplikasi Sistem Fuzzy Untuk Diagnosa Penyakit Jantung Koroner (Coronary Heart Disease). *Universitas Negeri Yogyakarta*.

Wihardi, Y. (2013). K-Folds Cross Validation. Retrieved April 25, 2018, from <http://blog.yayaw.web.id/riset/k-folds-cross-validation>

Wurdianarto, S. R., Novianto, S., & Rosyidah, U. (2014). Perbandingan Euclidean Distant dengan Canberra Distance Pada Fce Recognition, *13*(1), 31–37.

