

TUGAS AKHIR

ANALISIS PERBANDINGAN ALGORITMA NAIVE BAYES DAN K-NEAREST NEIGHBOR UNTUK KLASIFIKASI MULTI DATASET

Disusun Untuk Melengkapi dan Memenuhi Syarat Kelulusan Program Strata 1
Program Studi Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Jember



Diajukan oleh :

EKA BAGUS SUSANTO

NIM. 1410651097

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

2019

HALAMAN PENGESAHAN
ANALISIS PERBANDINGAN ALGORITMA NAIVE BAYES DAN K
NEIRS NEIGHBOR UNTUK KLASIFIKASI MULTI DATA SET

Oleh :

EKA BAGUS SUSANTO
1410651097

Telah Mempertanggung Jawabkan Tugas Akhirnya Pada Seminar Hasil Tugas Akhir Hari
Senin Tanggal 25 Maret 2019 Sebagai Salah Satu Syarat Kelulusan Dan Mendapatkan Gelar
Sarjana Komputer (S.Kom)
Di Program Studi Teknik informatika Fakultas Teknik
Universitas Muhammadiyah Jember

Disetujui Oleh :

Dosen penguji :
Penguji I

Dosen Pembimbing :
Pembimbing I

Wiwik Suharso, S.Kom., M.Kom
NIP.19760906 200501 1003

Triawan Adi Cahyanto, M.Kom
NPK.12 03 719

Penguji II

Pembimbing II

Ilham Saifudin S.Pd., M.Si
NPK .19891031 1 1703811

Reni Umilasari, S.Pd., M.Si
NPK.19910728 1 1703813

Mengesahkan , Dekan
Fakultas Teknik

Mengetahui,
Ketua Program Studi Teknik informatika

Ir. Subartinah, MT
NPK.95 05 246

Yeni Dwi Rahayu, S. ST., M. Kom.
NPK.11 03 590

SURAT PERYATAAN

Yang bertanda tangan dibawah ini ,saya :

Nama : Eka bagus susanto

NIM : 1410651097

Jurusan : Teknik informatika

Menyatakan dengan sesungguhnya bahwa karya ilmiah berupa Tugas Akhir yang berjudul : **“Analisis Perbandingan Algoritma Naive bayes dan K-Nearest Neighbor Untuk Klasifikasi Multi Data Set”** , adalah benar-benar karya sendiri ,kecuali jika dalam pengutipan substansi disebutkan sumbernya dan belum pernah diajukan pada intitusi manapun serta bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung timggi.

Demikian pernyataan ini saya buat dengan sebenar-benarnya,tanpa adanya tekanan dan paksaan dari pihak manapun serta bersedia mendapatkan sanksi akademik jika ternyata dikemudian hari pernyataan ini tidak benar.

Jember, 25 Maret 2019

Hormat kami,

Eka Bagus Susanto
Nim.1410651097

HALAMAN PERSEMBAHAN

Alhamdulillah kehadiran Allah SWT yang telah memberikan jalan-Nya sehingga tugas akhir ini berhasil diselesaikan .dalam penelitian yang dituliskan ini penulis mempersembahkan karya ini untu orang orang yang ikut membantu dalam usaha,pemikiran,inspirasi dan semangat dalam proses penyelesaian penelitian ini,antara lain adalah :

1. Allah SWT atas segala rahmat dan karuniaNYA dalam mempermudah segala urusan sayadalam rangka menyusun tugas akhir ini tanpa bantuanya saya tidak akan mampu menyelesaikan tugas akhir ini dengan sebaik-baiknya.
2. Kedua orang tua saya bapak sugianto dan ibu sri mubayamah ,terimakasih atas segala bantuan doa dan semangat dalam rangka penyelesaian tugas akhir ini.
3. kepada para sahabat Agil , Luki ,Arif , saudara sekaligus sahabat Roby terimakasih atas doa,bantuan dan segenap nasehat yang diberikan selama ini ,semoga Allah SWT membalasnya penuh barokah .
4. Teman-teman seperjuangan mahasiswa Universitas Muhammadiyah Jember dan Program Studi Informatika hingga saya mendapatkan gelar Sarjana komputer ini.
5. Terakhir almamaterku tercinta Universitas Muhammadiyah Jember dan Program Studi Tehnik Informatika hingga saya mendapat gelar Sarjana Komputer ini.

Akhirnya , dengan segala kerendahan hati penulis menyadari masih banyak terdapat kekurangan–kekurangan,sehingga penulis mengharapkan adanya saran dan kritik yang bersifat membangun demi kesempurnaan tugas akhir ini.

UNGKAPAN TERIMA KASIH

Alhamdulillah ,atas berkah dan karunia Allah SWT penulis dapat menyelesaikan Tugas Akhir ini dan dalam pelaksanaan pembuatannya penulis banyak mendapatkan bimbingan dan dukungan dari berbagai pihak .Oleh karena itu pada kesempatan kali ini Penulismenyampaikan terima kasih yang tulus dan ikhlas kepada :

1. **Allah SWT** ,karenanikmat .perlindungan ,Pertolongan, dan Ridho-Nya saya mampu menyelesaikan Tugas Akhir ini serta hambanya yang termulia Nabi besar Muhammad SAW.
2. **Bapak Sugianto ,Ibu Sri mubayamah,Adik dan keluarga** yang telah memberikan Doa,cinta dan kasih sayang baik itu berupa dorongan moril maupun spirituil.
3. **Ibu Yeni Dwi Rahayu , ST., M. M.Kom** kaprodi Tehnik Informatika Universitas Muhammadiyah Jember.
4. **Triawan Adi cahyanto M,Kom dan Reni umilasari S.Pd , M.Si** selaku dosen pembimbing banyak meluangkan waktu sampai tugas akhir ini selesai
5. Seluruh Dosen penguji : **Wiwik Suharso S.Kom , M.Kom dan Ilham Saifudin S.Pd, M.Si** yang bersedia meluangkan waktu untuk menguji dan mengoreksi hasil tugas akhir kami ini .
6. **Dosen dan Staf di fakultas teknik** ,yang telah memberikan ilmunya yang akan sangat berguna untuk penulis dan masyarakat umum.
7. Untuk teman teman **Kos ijo lumut** terima kasih telah membuat hidupku jauh lebih berwarna dan bermakna.
8. Seluruh **teman teman jurusan Tehnik Informatika dan khususnya angkatan 2014**,dan seluruh teman saya yang mengenal saya di Universitas Muhammadiyah Jember .

Semoga Allah SWT senantiasa memberikan perlindungan dan memberikan balasan yang lebih dikemudian hari.

Harapan kami sebagai penulis semoga dengan terselesainya Tugas Akhir in Dapat bermanfaat bagi kami khususnya untuk pengembanga ilmu pengetahuan dan Teknologi

Jember , 25 Maret 2019

Penulis

KATA PENGANTAR

نَمُوْا بِرَحْمَةِ رَبِّكَ الْعَلِيْمِ

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

ANALISIS PEBANDINGAN ALGORITMA NAIVE BAYES K – NEAREST NEIGHBORS UNTUK KLASIFIKASI MULTI DATA SET.

Dalam menyelesaikan Tugas Akhir ini ,kami berpegang pada teori yang pernah kami dapatkan dan bimbingan dari dosen pembimbing Tugas Akhir .Dan pihak-pihak lainyang sangat membantu hingga sampai terselesaikanya 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 ntuk para pembaca pada khususnya.

Jember ,25 Maret 2019

Penulis

MOTTO

Kerjakan Atau Mati

(Warren Buffett)

Jangan pernah jadi orang biasa karena itu membosankan.

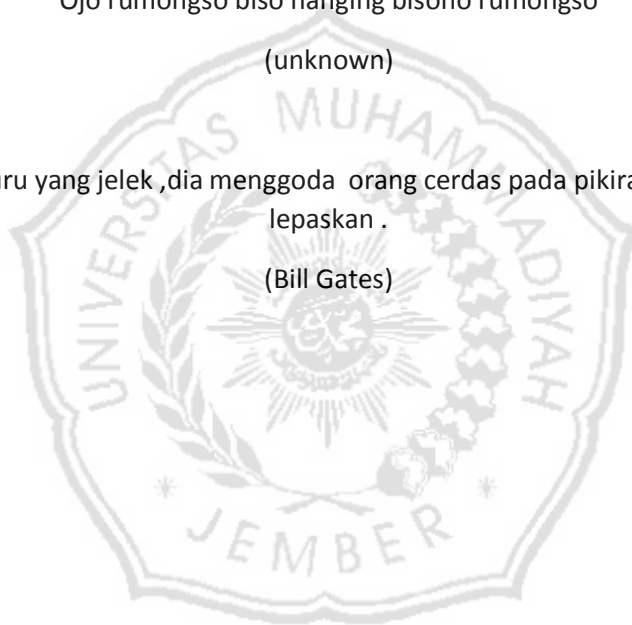
(Eka Bagus Susanto)

Ojo rumongso biso nanging bisoho rumongso

(unknown)

Kesuksesan adalah guru yang jelek ,dia menggoda orang cerdas pada pikiran yang mereka tak bisa lepaskan .

(Bill Gates)



DAFTAR ISI

BAB I PENDAHULUAN.....	7
1.1 Latar Belakang Masalah	8
1.2 Rumusan masalah	9
1.3 Tujuan	9
1.4 Batasan masalah.....	9
1.5 Manfaat	9
BAB II TINJAUAN PUSTAKA	10
2.1 Penelitian Terdahulu	10
2.2 Klasifikasi	11
2.1.1 Konsep Klasifikasi	11
2.1.2 Model	12
2.1.3 Pengukuran Kinerja Klasifikasi	14
2.3 Pengertian Analisis	17
2.4 Naïve Bayes	17
2.4.1 Teorema Bayes.....	17
2.4.2 Naïve Bayes untuk Klasifikasi	20
2.5 K-Nearest Neighbor	23
2.5.1 Pengertian K-Nearest Neighbor	23
2.5.2 Karakteristik Klasifikasi dengan Nearest Neighbor.....	25
2.5.3 Contoh Perhitungan Algoritma K-Nearest Neighbor.....	26
2.6 K-fold Cross Validation	29
2.7 Tool Weka.....	31
2.7.1 Pengertian Tool Weka.....	31
2.7.3 Kelebihan dan Kekurangan Tool Weeka	31
BAB III METODOLOGI PENELITIAN	34
3.1 Studi Literatur	35
3.2 Pengumpulan Data	35
3.3 Preprocessing data	35

3.3.1 Pembersihan Data	36
3.3.2 Analisis Data.....	37
3.4 Analisa Spesifikasi.....	38
3.5 Implementasi.....	39
3.6 Perhitungan Manual	40
3.6.1 Flowchart Naive bayes.....	40
3.6.2 Perhitungan Manual <i>Naïve</i> Bayes.....	41
3.6.3 Hasil perhitungan Naive bayes	43
3.6.4 Perhitungan K- Nearest Neighbor.....	43
3.6.5 Flowchart K- nearest Neighbor.....	43
3.6.6 Perhitungan Manual	44
3.7 Pengujian.....	44
BAB IV HASIL DAN PEMBAHASAN	47
4.1 Implementasi dan Pengujian Algoritma <i>Naïve</i> Bayes dalam Weka	47
4.2 Implementasi dan Pengujian Algoritma K-Nearest Neighbor dalam Weka	49
4.3 Analisis	50
BAB V KESIMPULAN DAN SARAN.....	58
5.1 Kesimpulan	58
5.2 Saran	58
DAFTARPUSTAKA.....	60
LAMPIRAN-LAMPIRAN	61

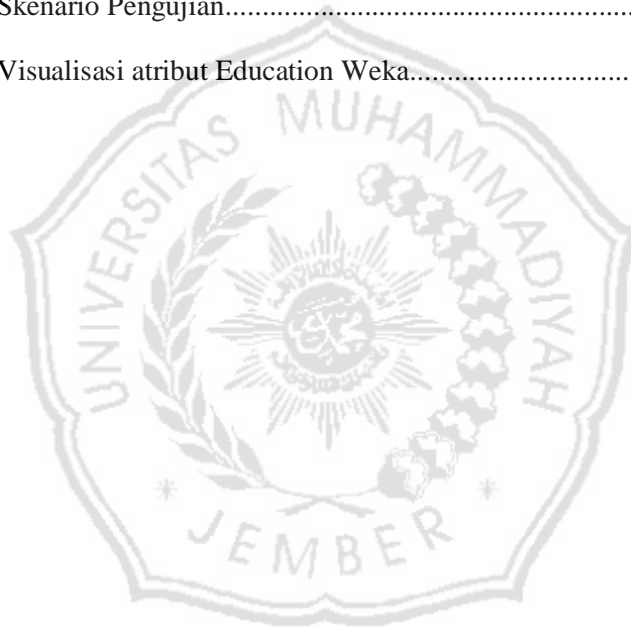
DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	7
Tabel 2.2 Matriks Konfusi untuk Klasifikasi dua kelas	12
Tabel 2.3 Parameter dan keterangan Naive Bayes	15
Tabel 2.4 Contoh data training.	24
Tabel 2.5 Tabel Jarak euclid (euclidean distance)	24
Tabel 2.6 Tabel Jarak terkecil (euclidean distance)	25
Tabel 2.7 Tabel Jarak terbaik (euclidean distance)	25
Tabel 2.8 Tabel Jarak euclid hasil terbaik (euclidean distance)	25
Tabel 2.9 Tabel data set awal.....	26
Tabel 2.10 Contoh Hasil Klasifikasi.....	26
Tabel 2.11 Contoh data set	27
Tabel 2.12 Contoh eksperimen	27
Tabel 3.1 Data set penelitian	32
Tabel 3.2 Data Training bank marketing sebelum di procesing data.....	33
Tabel 3.3 Data Training bank marketing setelah di procesing data.....	33
Tabel 3.4 Jumlah Data Hasil Pembersihan	33
Tabel 3.5 Kategori Jumlah record.....	34
Tabel 3.6 Kategori jumlah Variabel.....	34
Tabel 3.7 Kategori data set penelitian	35
Tabel 3.8 Spesifikasi Hardware dan Software.....	35
Tabel 3.9 Hasil Perhitungan Data Training Bank Marketing.....	40
Tabel 3.7 Data Testing	40
Tabel 3.7 Hasil Perhitungan K-Nearest Neighbor.....	41

Tabel 4.1 Hasil Pengujian Data 1	44
Tabel 4.2 Hasil Pengujian Data 2	45
Tabel 4.3 Hasil Pengujian Data 1	46
Tabel 4.4 Hasil Pengujian Data 2	47
Tabel 4.5 Tabel Analisa Akurasi Pada Teknik Pengujian Cross Validation K = 2	48
Tabel 4.6 Tabel Analisa Akurasi Pada Teknik Pengujian Cross Validation K = 3	48
Tabel 4.7 Tabel Analisa Akurasi Pada Teknik Pengujian Cross Validation K = 4.....	48
Tabel 4.8 Tabel Analisa Akurasi Pada Teknik Pengujian Cross Validation K = 5.....	48
Tabel 4.9 Tabel Analisa Akurasi Pada Teknik Pengujian Cross Validation K = 6.....	48
Tabel 4.10 Tabel Analisa Akurasi Pada Teknik Pengujian Cross Validation K = 7.....	49
Tabel 4.11 Tabel Analisa Akurasi Pada Teknik Pengujian Cross Validation K = 8.....	49
Tabel 4.12 Tabel Analisa Akurasi Pada Teknik Pengujian Cross Validation K = 9.....	49
Tabel 4.13 Tabel Analisa Akurasi Pada Teknik Pengujian Cross Validation K = 10.....	49
Tabel 4.14 Tabel Analisa Akurasi Pada Teknik Pengujian Cross Validation K = UTS..	49
Tabel 4.15 Perbandingan dua dataset	50
Tabel 4.16 Kesimpulan Perbandingan Algoritma Naive Bayes & K-Nearest Neighbor.	52

DAFTAR GAMBAR

Gambar 2.1 K- Nearest Neighbor dengan nilai K yang besar.....	20
Gambar 3.1 Alur Pengujian dengan Weka	36
Gambar 3.2 Flowchart Alur Naive bayes.....	37
Gambar 3.3 Alur Perhitungan K -Nearest Neighbor.....	40
Gambar 3.4 Skenario Pengujian.....	42
Gambar 4.1 Visualisasi atribut Education Weka.....	52



DAFTAR PUSTAKA

Prasetyo,Eko (2012). Data Mining konsep dan aplikasi menggunakan matlab.yogyakarta:Andi.

Dwi Luki arisandy(2017) Analisa perbandingan algoritma dan algoritma c4.5 untuk klasifikasi multi data

Nur Khotimah dan Deden Istiawan (2018) berjudul “Perbandingan Algoritma C4.5, Naïve Bayes dan K-Nearest Neighbour untuk Prediksi Lahan Kritis di Kabupaten Pematang Siantar

Wihardi yaya .(2013).“K-foldCross Validation”.Diakses dari [http://blog.yayaw.web.id/riset/k-fold-cross-validation] .

<http://repository.telkomuniversity.ac.id/pustaka/94560/analisis-perbandingan-metode-k-nearest-neighbor-dan-naive-bayes-classifier-dalam-klasifikasi-teks.html>

<http://www.idx.co.id/id-id/beranda/perusahaantercatat/profilperusahaantercatat.aspx>

<http://www.cs.waikato.ac.nz/ml/people.html>

putubuku, “Recall & Precision,” *Ilmu Perpustakaan & Informasi – diskusi dan ulasanringkas*,27-Mar-

2008.[Online].Available: <http://iperpin.wordpress.com/2008/03/27/recall-precision/>. [Accessed: 16-Jun-2013]