

TUGAS AKHIR

**KLASIFIKASI *E-MAIL SPAM* MENGGUNAKAN
METODE *K-NEAREST NEIGHBOR***



ALAN FITRIYANTO

14 1065 1086

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH JEMBER

2019

HALAMAN PENGESAHAN

**KLASIFIKASI E-MAIL SPAM MENGGUNAKAN
METODE K-NEAREST NEIGHBOR**

Oleh:

ALAN FITRIYANTO

14 1065 1086

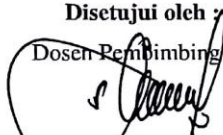
Telah mempertanggung jawabkan Proposal Tugas Akhirnya pada sidang Tugas Akhir tanggal 19 Agustus 2019 sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom).

Di

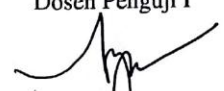
Universitas Muhammadiyah Jember

Disetujui oleh :

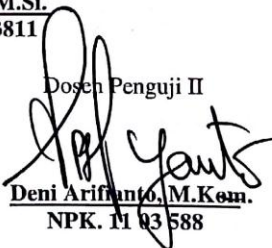
Dosen Pembimbing


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

Dosen Penguji I


Agung Nilogiri, S.T., M.Kom.
NIP. 19770330 200501 1 002

Dosen Penguji II


Deni Arifianto, M.Kom.
NPK. 11 03 588

Mengesahkan,
Dekan Fakultas Teknik Universitas
Muhammadiyah Jember


W. Suhartinah, M.T.
NPK. 95 05 246

Mengetahui,
Ketua Program Studi Teknik Informatika


Seti Dwi Rahayu, M.Kom.
NPK. 11 03 590

PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : ALAN FITRIYANTO

Nim : 14 1065 1086

Menyatakan dengan sesungguhnya bahwa karya tulis ilmiah yang berjudul “Klasifikasi *E-Mail Spam* Menggunakan Metode *K-Nearest Neighbor*” adalah benar-benar hasil karya sendiri, kecuali jika disebutkan sumbernya dan belum pernah diajukan pada institusi mana pun, serta bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa adanya tekanan dan paksaan dari pihak mana pun serta bersedia mendapat sanksi akademik jika ternyata di kemudian hari pernyataan ini tidak benar.

Jember, 26 Maret 2019

menyatakan,

FITRIYANTO

NIM. 14 1065 1086

KATA PENGANTAR

Puji syukur saya panjatkan kepada Allah SWT karena rahmat dan karuniaNya-lah saya dapat menyelesaikan tugas akhir ini dengan baik yang berjudul “Klasifikasi E-Mail Spam Menggunakan Metode K-Nearest Neighbor” disusun sebagai salah satu syarat untuk mengerjakan skripsi pada program S1 di Fakultas Teknik Prodi Teknik Informatika Universitas Muhammadiyah Jember.

Saya menyadari dalam penyusunan proposal skripsi ini tidak selesai tanpa bantuan dari berbagai pihak. Karena itu pada kesempatan ini saya ingin mengucapkan terima kasih kepada :

1. Ibu Yeni Dwi Rahayu M.Kom selaku Kepala Program Studi Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember.
2. Bapak Ilham Saifudin, S.Pd.,M.Si selaku pembimbing yang telah meluangkan waktunya untuk membimbing saya terhadap penyelesaian pembuatan proposal Tugas Akhir ini.
3. Orang tua, saudara-saudara serta teman-teman saya atas doa, bimbingan dan kasih sayang yang selalu tercurah selama ini.
4. Kepada pihak-pihak yang telah membantu saya sehingga tidak dapat saya sebutkan satu persatu.

Saya menyadari bahwa proposal tugas akhir ini jauh dari sempurna, oleh karena itu saya mengharapkan saran dan kritik yang bersifat membangun dari segenap pembaca. Akhirnya, semoga proposal tugas akhir ini dapat berguna dan bermanfaat terutama bagi pihak-pihak yang tertarik untuk mengkaji dan mengembangkannya.

Jember, 30 September 2019

Penulis

Daftar Isi

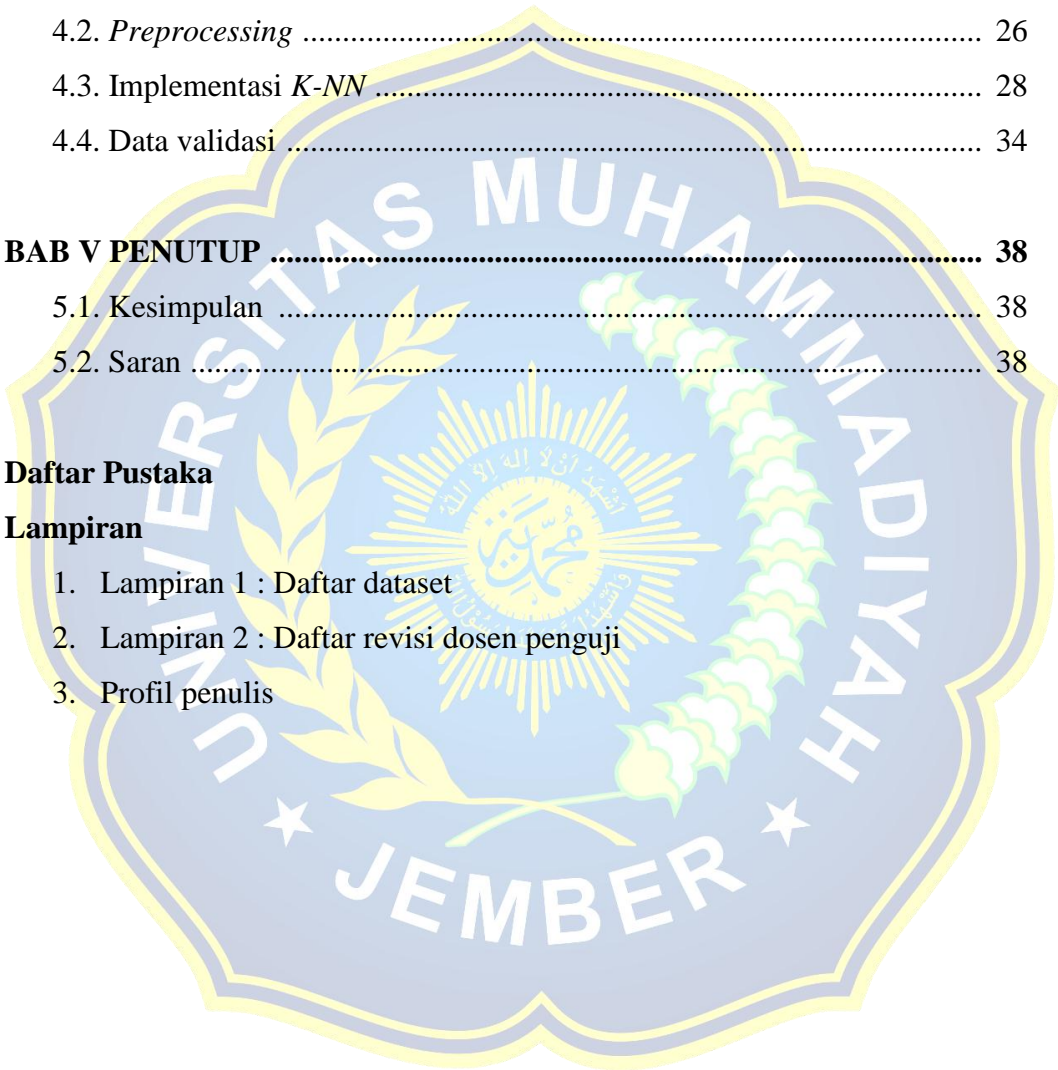
Halaman Judul	i
Halaman Persetujuan	ii
Halaman Pengesahan	iii
Halaman Keaslian	iv
Abstrak	v
Abstract	vi
Halaman Motto	vii
Halaman Persembahan	viii
Kata Pengantar	ix
Daftar Isi	x
Daftar Gambar	xii
Daftar Tabel	xiii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan	2
1.4. Manfaat	3
1.5. Batasan Masalah	3
BAB II KAJIAN PUSTAKA	5
2.1. Penelitian Terkait	5
2.2. <i>Email</i>	7
2.3. <i>Email Spam</i>	8
2.4. <i>Data Mining</i>	10
2.5. Klasifikasi	10
2.6. <i>K Nearest Neighbor</i>	11
2.7. <i>Confusion Matrix</i>	13

BAB III METODE PENELITIAN	15
3.1. Tahapan Penelitian	15
3.2. Contoh Kasus	18
BAB IV IMPLEMENTASI DAN PENGUJIAN	24
4.1. Deskripsi data	24
4.2. <i>Preprocessing</i>	26
4.3. Implementasi <i>K-NN</i>	28
4.4. Data validasi	34
BAB V PENUTUP	38
5.1. Kesimpulan	38
5.2. Saran	38

Daftar Pustaka

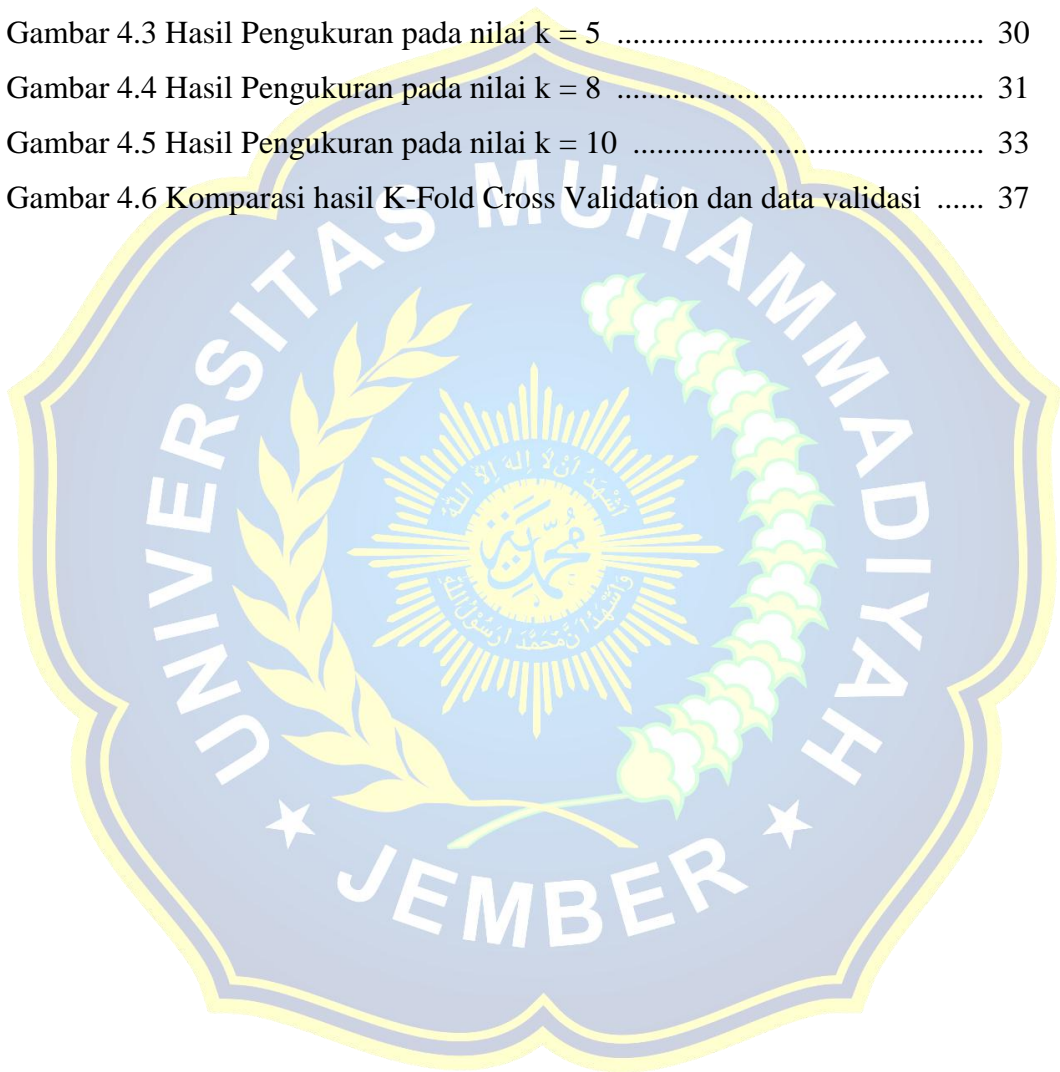
Lampiran

1. Lampiran 1 : Daftar dataset
2. Lampiran 2 : Daftar revisi dosen penguji
3. Profil penulis



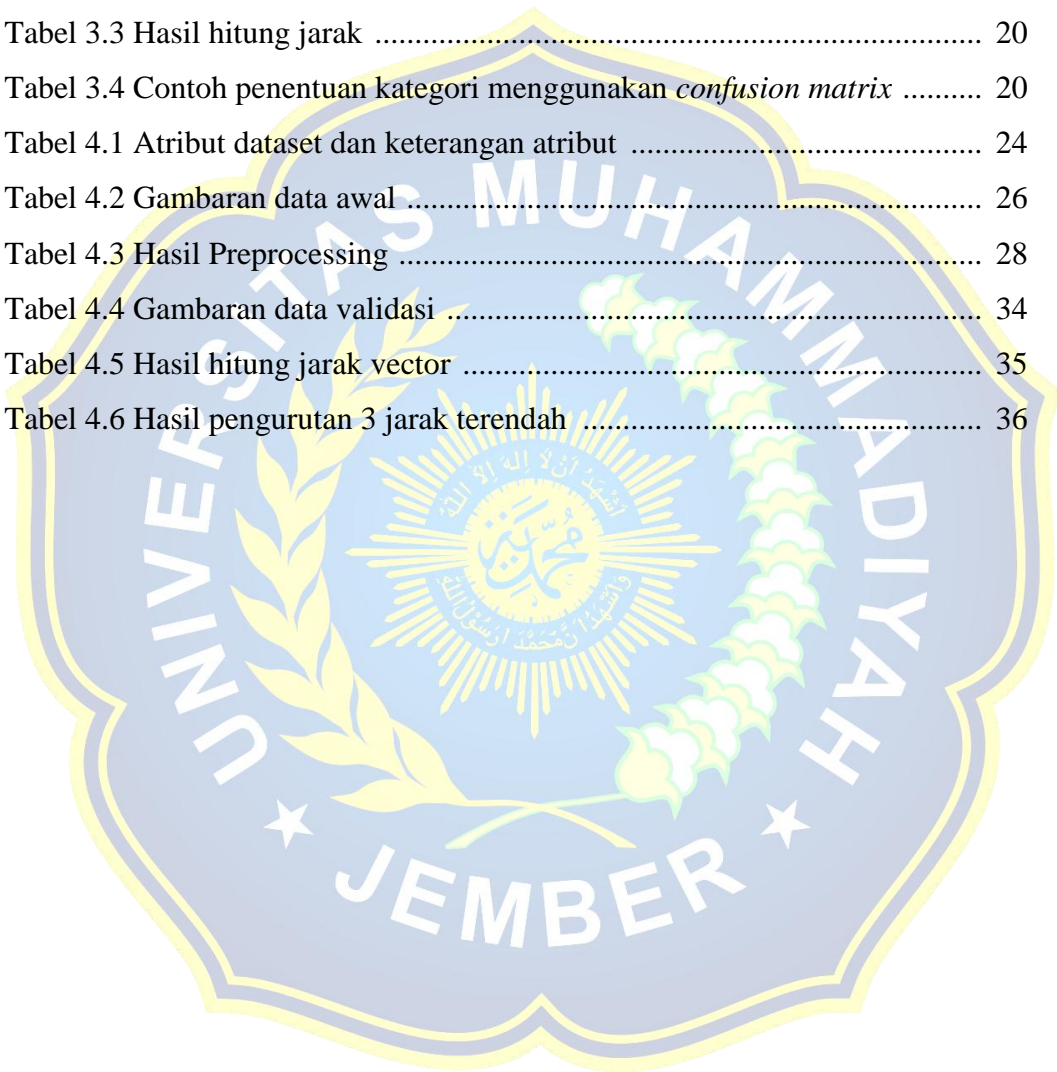
Daftar Gambar

Gambar 2.1 Ilustrasi Solusi pada KNN (Sumber : Kursini, at all. 2009)	12
Gambar 3.1 Tahapan Penelitian	15
Gambar 4.1 Hasil Pengukuran pada nilai $k = 2$	29
Gambar 4.2 Hasil Pengukuran pada nilai $k = 2$	29
Gambar 4.3 Hasil Pengukuran pada nilai $k = 5$	30
Gambar 4.4 Hasil Pengukuran pada nilai $k = 8$	31
Gambar 4.5 Hasil Pengukuran pada nilai $k = 10$	33
Gambar 4.6 Komparasi hasil K-Fold Cross Validation dan data validasi	37



Daftar Tabel

Tabel 2.1 Penelitian terkait	5
Tabel 2.2 Penetapan kriteria dengan <i>Confusion matrix</i>	14
Tabel 3.1 Contoh data latih	18
Tabel 3.2 Contoh data uji	19
Tabel 3.3 Hasil hitung jarak	20
Tabel 3.4 Contoh penentuan kategori menggunakan <i>confusion matrix</i>	20
Tabel 4.1 Atribut dataset dan keterangan atribut	24
Tabel 4.2 Gambaran data awal	26
Tabel 4.3 Hasil Preprocessing	28
Tabel 4.4 Gambaran data validasi	34
Tabel 4.5 Hasil hitung jarak vector	35
Tabel 4.6 Hasil pengurutan 3 jarak terendah	36



Lampiran 1

Daftar dataset

Untuk mempermudah pembaca, dataset dapat diakses pada laman-laman berikut mengingat jumlah dataset 4601 record dengan jumlah atribut 59, maka penulis akan mencantumkan laman untuk akses pembaca.

Data untuk informasi atribut dataset

<https://archive.ics.uci.edu/ml/datasets/Spambase>

Dataset untuk download

<https://archive.ics.uci.edu/ml/machine-learning-databases/spambase/>



Lampiran 2



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

Jl. Karimata 49 Telp. (0331) 336728 Fax. (0331) 337957 Kotak Pos 104 Jember 68121

**DAFTAR REVISI PENGUJI 1
SIDANG TUGAS AKHIR**

Nama Mahasiswa : ALAN FITRIYANTO
Nomor Induk Mahasiswa : 1410651086
Judul Tugas Akhir : Klasifikasi email spam menggunakan metode K Nearest Neighbor
Hari / Tanggal : Jumat / 16 Agustus 2019
Jam : 09:00 WIB
Tempat : Cc.2.2

Bab/Halaman	Uraian	Keterangan
	kutipasi sumber	
	penjabaran K-NN ke long	
	Bahan masalah data HP 1999 & k	
	Penjabaran atribut	
	Pre processing	

Dosen Penguji 1

DENI ARIFANTO, M.Kom.

NB : Untuk Mahasiswa



PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER

Jl. Karimata 49 Telp. (0331) 336728 Fax. (0331) 337957 Kotak Pos 104 Jember 68121

DAFTAR REVISI PENGUJI 2
SIDANG TUGAS AKHIR

Nama Mahasiswa : ALAN FITRIYANTO
Nomor Induk Mahasiswa : 1410651086
Judul Tugas Akhir : Klasifikasi email spam menggunakan metode K Nearest Neighbor
Hari / Tanggal : Jumat / 16 Agustus 2019
Jam : 09:00 WIB
Tempat : Cc.2.2

Bab/Halaman	Uraian	Keterangan
	penelitian kata? asyq diperbaiki	
	manfaat diperbaiki	
	tabel penelitian terdahulu diperbaiki	
	tabel confusion matrix diperbaiki	
	recall & error rate digunakan?	
	cek ulang hasil pengujian dan ringkasan	
	daftar pustaka diperagakan	

Dosen Penguji 2

AGUNG NILOGURI, S.T., M.Kom.

Lampiran 3 Profil Penulis

Profil Penulis



Nama : Alan Fitriyanto
NIM : 14 1065 1086
Tempat, tanggal lahir : Bondowoso, 21 Februari 1996
Alamat : Jl. A.Yani Gg. 3 RT.03 RW.01 Kel. Dawuhan
Kab. Situbondo
Email : alanfitriyanto2@gmail.com

Riwayat Pendidikan

- SD Negeri 6 Dawuhan Lulus Tahun 2008
- Mts Negeri 1 Purwoasri Lulus Tahun 2011
- SMK Negeri 1 Situbondo Lulus Tahun 2014