

TUGAS AKHIR

ANALISIS PEMODELAN METODE *K-NEAREST NEIGHBOR* TERHADAP KLASIFIKASI KESIAPAN SISWA KELAS IX SMP AL GAZALI DALAM MENGHADAPI UNBK



Disusun Oleh :

FAHRIADI

14 1065 1018

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH JEMBER

2020

TUGAS AKHIR

**ANALISIS PEMODELAN METODE *K-NEAREST NEIGHBOR*
TERHADAP KLASIFIKASI KESIAPAN SISWA KELAS IX
SMP AL GAZALI DALAM MENGHADAPI UNBK**



Oleh :

FAHRIADI

NIM. 14 1065 1018

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

2020

**ANALISIS PEMODELAN METODE *K-NEAREST NEIGHBOR*
TERHADAP KLASIFIKASI KESIAPAN SISWA KELAS IX
SMP AL GAZALI DALAM MENGHADAPI UNBK**

Tugas Akhir

**Ditulis dan diajukan untuk memenuhi persyaratan mendapatkan gelar
Sarjana Komputer pada Program Studi Teknik Informatika Fakultas
Teknik Universitas Muhammadiyah Jember**



Oleh:

**FAHRIADI
14 1065 1018**

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

2020

**ANALISIS PEMODELAN METODE *K-NEAREST NEIGHBOR*
TERHADAP KLASIFIKASI KESIAPAN SISWA KELAS IX
SMP AL GAZALI DALAM MENGHADAPI UNBK**

Diajukan oleh :

**FAHRIADI
14 1065 1018**

Tugas Akhir ini telah disetujui untuk dipertahankan di hadapan Tim Penguji Skripsi Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember.

Persetujuan Pembimbing

Dosen Pembimbing 1

Dosen Pembimbing 2



Dewi Lusiana,Ir.MT
NPK. 04 10 624



Daryanto,S.Kom,M.Kom
NPK. 11 03 589

HALAMAN PENGESAHAN

ANALISIS PEMODELAN METODE *K-NEAREST NEIGHBOR* TERHADAP KLASIFIKASI KESIAPAN SISWA KELAS IX SMP AL GAZALI DALAM MENGHADAPI UNBK


Oleh :
FAHRIADI
14 1065 1018

Telah mempertanggung jawabkan Laporan Tugas Akhirnya pada sidang Tugas Akhir tanggal 09 September 2020 sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom).


Di
Universitas Muhammadiyah Jember

Disetujui oleh,

Dosen Penguji I


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

Dosen Penguji II


Deni Arifanto, M.Kom
NPK : 11 08 588


Mengesahkan,
Dekan Fakultas Teknik
Universitas Muhammadiyah Jember


Nanang Saiful Rizal, S.T.,MT
NPK : 09 03 315


Dosen Pembimbing I


Dewi Lusiana, Ir.MT
NPK : 04 10 624

Dosen Pembimbing II


Daryanto, S.Kom.M.Kom
NPK : 11 03 589

Mengetahui,
Ketua Program Studi Teknik
Informatika


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

PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : FAHRIADI

Nim : 14 1065 1018

Menyatakan dengan sesungguhnya bahwa karya tulis ilmiah yang berjudul **“ANALISIS PEMODELAN METODE *K-NEAREST NEIGHBOR* TERHADAP KLASIFIKASI KESIAPAN SISWA KELAS IX SMP AL GAZALI DALAM MENGHADAPI UNBK”** 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, 22 Oktober 2020



FAHRIADI
NIM 14 1065 1018

MOTTO

“Jadikanlah sabar dan sholat sebagai penolongmu, dan sesungguhnya yang demikian itu adalah berat, kecuali bagi orang-orang yang khusyu”

(Qs. Al Baqoroh:45)

“ sesungguhnya sesudah kesulitan itu ada kemudahan, maka apabila telah selesai (urusan dunia) bersungguh- sungguhlah (dalam beribadah). Dan hanya kepada Tuhan-Mu lah ber harap”

(Qs. Al Insyirah: 6-8)



PERSEMBAHAN

Alhamdulillah ku panjatkan kepada Allah SWT, atas segala rahmat dan juga kesempatan dalam menyelesaikan tugas akhir skripsi saya dengan segala kekurangannya. Segala syukur ku ucapkan kepadaMu Ya Rabb, karena sudah menghadirkan orang-orang berarti disekeliling saya. Yang selalu memberi semangat dan doa, sehingga skripsi saya ini dapat diselesaikan dengan baik.

Untuk karya yang sederhana ini, maka saya persembahkan untuk

1. Ayahanda dan Ibunda tercinta dan tersayang

Apa yang saya dapatkan hari ini, belum mampu membayar semua kebaikan, keringat, dan juga air mata bagi saya. Terima kasih atas segala dukungan kalian, baik dalam bentuk materi maupun moril. Karya ini saya persembahkan untuk kalian, sebagai wujud rasa terima kasih atas pengorbanan dan jerih payah kalian sehingga saya dapat menggapai cita-cita. Kelak cita-cita saya ini akan menjadi persembahan yang paling mulia untuk Ayah dan Ibu, dan semoga dapat membahagiakan kalian.

2. Adik tercinta

Untuk Adikku Ahmad Fahrul Reza, tiada waktu yang paling berharga dalam hidup selain menghabiskan waktu denganmu. Walaupun saat dekat kita sering bertengkar, tapi saat jauh kita saling merindukan.

3. Dosen Pembimbing dan Dosen Penguji

Kepada Ibu Dewi Lusiana, Ir.MT dan bapak Daryanto, S.Kom, M.Kom, selaku dosen pembimbing serta bapak Agung Nilogiri S.T., M.Kom dan bapak Deni Arifianto, M.Kom, selaku dosen penguji terima kasih karena sudah menjadi orang tua kedua saya di Kampus.

4. Sahabat dan seluruh teman di kampus tercinta

Tanpa kalian mungkin masa-masa kuliah saya akan menjadi biasa-biasa saja, maaf jika banyak salah dengan maaf yang tak terucap.

KATA PENGANTAR

Puji syukur ke hadirat Allah Swt. atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan tugas akhir yang berjudul “**ANALISIS PEMODELAN METODE *K-NEAREST NEIGHBOR* TERHADAP KLASIFIKASI KESIAPAN SISWA KELAS IX SMP AL GAZALI DALAM MENGHADAPI UNBK**”. Tugas akhir ini disusun untuk memenuhi salah satu syarat dalam menyelesaikan pendidikan strata satu (S1) pada Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember.

Penyusunan tugas akhir ini tidak lepas dari bantuan berbagai pihak. Oleh karena itu, penulis menyampaikan ucapan terima kasih kepada:

1. Dr. Hanafi M,pd. selaku Rektor Universitas Muhammadiyah Jember.
2. Dr. Nanang Saiful Rizal, M.T., selaku Dekan Fakultas Teknik.
3. Triawan Adi Cahyanto, M.Kom, selaku Ketua Jurusan Teknik Informatika.
4. Dewi Lusiana, Ir.MT, selaku Pembimbing I yang telah meluangkan waktu, memotivasi, membimbing, dan memberi masukan demi sempurnanya tugas akhir ini.
5. Daryanto, S.Kom, M.Kom, selaku Pembimbing II yang telah meluangkan waktu, memotivasi, membimbing, dan memberi masukan demi sempurnanya tugas akhir ini.
6. Agung Nilogiri, S.T., M.Kom, selaku Dosen Penguji I yang telah meluangkan waktu dan memberikan masukan demi sempurnanya tugas akhir ini.
7. Deni Arifianto, M.Kom, selaku Dosen Penguji II yang telah meluangkan waktu dan memberikan masukan demi sempurnanya tugas akhir ini.
8. Bapak dan Ibu Dosen Teknik Informatika Universitas Muhammadiyah Jember yang telah memberikan ilmu pengetahuan kepada penulis.
9. Staf Jurusan Teknik Informatika yang telah membantu dan memberikan petunjuk kepada penulis dalam proses menyelesaikan administrasi tugas akhir.
10. Karyawan UPT Perpustakaan Universitas Muhammadiyah Jember yang telah membantu penulis dalam mencari buku referensi.

11. Karyawan UPT Perpustakaan Fakultas Teknik Universitas Muhammadiyah Jember yang telah membantu penulis dalam mencari buku referensi.
12. Semua teman-teman Teknik Informatika dan Angkatan 2014 yang tidak bisa saya sebutkan satu per satu, yang setia membantu, dan memberikan inspirasi selama studi di Teknik Informatika.

Penulis akan menerima segala kritik dan saran dari semua pihak yang bersifat membangun. Semoga tugas akhir ini dapat bermanfaat.



**ANALISIS PEMODELAN METODE *K-NEAREST NEIGHBOR*
TERHADAP KLASIFIKASI KESIAPAN SISWA KELAS IX SMP AL
GAZALI DALAM MENGHADAPI UNBK**

FAHRIADI¹, Dewi Lusiana², Daryanto³

Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember
Fahrygundez@gmail.com, dewilusiana@unmuhjember.ac.id,
daryanto@unmuhjember.ac.id

ABSTRAK

SMP Al Gazali Jember adalah salah satu sekolah yang mempersiapkan siswanya menghadapi UNBK dengan cara melakukan kerja sama pengadaan *tryout* dengan lembaga bimbingan. Pada penelitian ini dilakukan analisis klasifikasi pada implementasi *machine learning*. Atribut yang digunakan dalam penelitian ini adalah nilai rapor tengah semester 1, rapor semester 1 dan empat kali nilai *tryout*. Nilai yang terbagi menjadi 4 atribut sesuai mata pelajaran ujian nasional yaitu Bahasa Indonesia, Matematika, Bahasa Inggris dan Ilmu Pengetahuan Alam dan klasifikasi siswa yaitu “siap” dan “belum” dalam menghadapi ujian nasional. Pada penelitian ini metode klasifikasi yang digunakan adalah *K Nearest Neighbor* dengan pengukuran jarak menggunakan *Euclidean Distance* dan nilai tetangga terdekat yaitu 3, 5, 7 dan 9. Pada penelitian ini juga digunakan skenario uji *cross fold validation* dengan nilai *fold* yaitu 2, 4, 5, 7 dan 10. Nilai akurasi tertinggi yang diperoleh metode *K Nearest Neighbor* adalah 100% yaitu pada *fold* 2 skenario 2, *fold* 4 skenario 1, 3 dan 4, *fold* 5 skenario 1, 4, dan 5, *fold* 7 skenario 1, 4, 5, 6, 7, *fold* 10 skenario 1, 2, 4, 6, 7, 8, 9 dan 10. Nilai presisi tertinggi yang diperoleh metode *K Nearest Neighbor* adalah 100% yaitu diperoleh oleh semua skenario uji dan metode *K Nearest Neighbor* dapat digunakan untuk mengetahui klasifikasi kesiapan siswa dalam menghadapi ujian nasional.

Kata kunci : Analisis, Klasifikasi, *K-Nearest Neighbor*, Siswa.

ANALYSIS OF MODELING THE K-NEAREST NEIGHBOR METHOD ON THE CLASSIFICATION OF READINESS CLASS IX STUDENTS OF AL GAZALI IN FACING UNBK

FAHRIADI¹, Dewi Lusiana², Daryanto³

Department of Informatics, Faculty of Engineering, University of Muhammadiyah Jember

Fahrygundez@gmail.com, dewilusiana@unmuhjember.ac.id,

daryanto@unmuhjember.ac.id

ABSTRACT

SMP Al Gazali Jember is one of the schools that prepares students for UNBK by cooperating with the procurement of tryouts with tutoring institutions. In this study, a classification analysis was carried out on the implementation of machine learning. The attributes used in this study were the first semester report cards, first semester report cards and four times the tryout values. The scores are divided into 4 attributes according to the national exam subjects, namely Indonesian, Mathematics, English and Natural Sciences and the classification of students, namely "ready" and "not yet" in facing the national exam. In this study the classification method used is K Nearest Neighbor with distance measurement using Euclidean Distance and the closest neighbor values, namely 3, 5, 7 and 9. In this study also used the cross fold validation test scenario with fold values of 2, 4, 5 7 and 10. The highest accuracy value obtained by the K Nearest Neighbor method is 100%, namely in fold 2 scenarios 2, fold 4 scenarios 1, 3 and 4, fold 5 scenarios 1, 4, and 5, fold 7 scenarios 1, 4, 5, 6, 7, fold 10 scenarios 1, 2, 4, 6, 7, 8, 9 and 10. The highest precision value obtained by the K Nearest Neighbor method is 100%, which is obtained by all test scenarios and the K Nearest Neighbor method can be used to know the classification of student readiness in facing the national exam.

Keywords: Analysis, Classification, K-Nearest Neighbor, Students.

Daftar Isi

Halaman Pengajuan	i
Halaman Persetujuan	ii
Halaman Pengesahan	iii
Halaman Pernyataan	iv
Halaman Motto	v
Halaman Persembahan	vi
Kata Pengantar	vii
Abstrak	viii
<i>Abstract</i>	ix
Daftar Isi	xi
Daftar Gambar	xiii
Daftar Tabel	xiv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Tujuan	3
1.4. Manfaat	3
1.5. Batasan Masalah	3
BAB II KAJIAN PUSTAKA	4
2.1. Penelitian Terkait	4
2.2. Prestasi Belajar	6
2.3. Data Mining	7
2.3.1. Tahapan Data Mining	8
2.3.2. Pengelompokan Data Mining	10
2.4. <i>K-Nearest Neighbor</i>	12
2.5. <i>Confusion Matrix</i>	13
2.6. <i>K-Fold Cross Validation</i>	15
2.7. <i>Orange Data Analytics</i>	16

2.8. <i>Imbalance Class</i>	16
2.9. <i>Synthetic Minority Oversampling Technique (SMOTE)</i>	17
2.10. <i>Waikato Environment for Knowledge Analysis (WEKA)</i>	18
BAB III METODE PENELITIAN	20
3.1. Tahapan Penelitian	20
3.2. Contoh Kasus	26
BAB IV IMPLEMENTASI DAN HASIL	31
4.1. Gambaran Data	31
4.2. <i>Preprocessing</i>	32
4.3. Implementasi <i>K Nearest Neighbor</i>	38
4.3.1. <i>Cross Fold Validation</i>	38
4.3.2. Data Validasi	59
BAB V PENUTUP	65
5.1. Kesimpulan	65
5.2. Saran	65
Daftar Pustaka	
Lampiran :	
1. Lampiran 1 : Data Cross Fold Validation (Hasil SMOTE)	
2. Lampiran 2 : Data Validasi	
3. Lampiran 3 : Sampel hitung data validasi	
4. Lampiran 4 : Confusion matrix data validasi	

Daftar Gambar

Gambar 2.1 Tahapan Data Mining	9
Gambar 2.2 Confusion Matrix 3 classes	14
Gambar 2.3 K-Fold Cross Validation 10 Fold	15
Gambar 3.1 Tahapan Penelitian	20
Gambar 3.2 Cross Validation k = 2	21
Gambar 3.3 Cross Validation k = 4	22
Gambar 3.4 Cross Validation k = 5	22
Gambar 3.5 Cross Validation k = 7	23
Gambar 3.6 Cross Validation k = 10	24
Gambar 4.1 Grafik kondisi siswa berdasarkan tiap mata pelajaran	32
Gambar 4.2 Partisi data Validasi dan skenario uji k Fold Cross Validation ...	35
Gambar 4.3 Proses SMOTE pada WEKA	36
Gambar 4.4 Perbandingan data sebelum SMOTE (kiri) dan setelah SMOTE (kanan)	37
Gambar 4.3 Perbandingan output pada data validasi	62

Daftar Tabel

Tabel 2.1 Penelitian Terkait	4
Tabel 2.2 Skenario partisi data latih dan uji pada Cross Fold Validation dengan total data 100	15
Tabel 3.1 Contoh pengukuran hasil klasifikasi tiap k pada skenario 2 Fold ..	21
Tabel 3.2 Contoh pengukuran hasil klasifikasi tiap k pada skenario 4 Fold ..	22
Tabel 3.3 Contoh pengukuran hasil klasifikasi tiap k pada skenario 5 Fold ..	23
Tabel 3.4 Contoh pengukuran hasil klasifikasi tiap k pada skenario 5 Fold ..	23
Tabel 3.5 Contoh pengukuran hasil klasifikasi tiap k pada skenario 10 Fold	24
Tabel 3.6 Contoh data latih	26
Tabel 3.7 Contoh data uji	26
Tabel 3.8 Contoh hasil perhitungan Distance	27
Tabel 3.9 Contoh hasil sorting 3 tetangga terdekat	27
Tabel 3.10 Contoh hasil klasifikasi K-NN	28
Tabel 3.11 Contoh data untuk SMOTE	28
Tabel 3.12 Contoh data yang ingin direplikasi	29
Tabel 3.13 Hasil pengukuran jarak untuk pengolahan SMOTE	29
Tabel 3.14 Hasil replikasi data	30
Tabel 4.1 Atribut dan sub atribut pada data awal	33
Tabel 4.2 Potongan data dan atribut awal pada data siswa	33
Tabel 4.3 Potongan data siswa setelah akumulatif	34
Tabel 4.4 Potongan data hasil proses SMOTE	37
Tabel 4.5 Perbandingan data awal dan setelah dilakukan SMOTE	38
Tabel 4.6 Potongan data untuk skenario uji Cross Validation	39
Tabel 4.7 Confusion Matrix pada Fold 2 skenario 1	40
Tabel 4.8 Confusion Matrix pada Fold 2 skenario 2	40
Tabel 4.9 Potongan data latih skenario 2.2	41
Tabel 4.10 Sample data uji skenario 2.2	42
Tabel 4.11 Hasil sample hitung Nearest Neighbor k = 3 pada skenario 2.2 ...	42
Tabel 4.12 Hasil perbandingan klasifikasi orange dan hitung manual	43

Tabel 4.13 Confusion Matrix pada Fold 4 skenario 1	43
Tabel 4.14 Confusion Matrix pada Fold 4 skenario 2	44
Tabel 4.15 Confusion Matrix pada Fold 4 skenario 3	44
Tabel 4.16 Confusion Matrix pada Fold 4 skenario 4	45
Tabel 4.17 Confusion Matrix pada Fold 5 skenario 1	46
Tabel 4.18 Confusion Matrix pada Fold 5 skenario 2	46
Tabel 4.19 Confusion Matrix pada Fold 5 skenario 3	47
Tabel 4.20 Confusion Matrix pada Fold 5 skenario 4	48
Tabel 4.21 Confusion Matrix pada Fold 5 skenario 5	48
Tabel 4.22 Confusion Matrix pada Fold 7 skenario 1	49
Tabel 4.23 Confusion Matrix pada Fold 7 skenario 2	49
Tabel 4.24 Confusion Matrix pada Fold 7 skenario 3	50
Tabel 4.25 Confusion Matrix pada Fold 7 skenario 4	51
Tabel 4.26 Confusion Matrix pada Fold 7 skenario 5	51
Tabel 4.27 Confusion Matrix pada Fold 7 skenario 6	52
Tabel 4.28 Confusion Matrix pada Fold 7 skenario 7	52
Tabel 4.29 Confusion Matrix pada Fold 10 skenario 1	53
Tabel 4.30 Confusion Matrix pada Fold 10 skenario 2	54
Tabel 4.31 Confusion Matrix pada Fold 10 skenario 3	54
Tabel 4.32 Confusion Matrix pada Fold 10 skenario 4	55
Tabel 4.33 Confusion Matrix pada Fold 10 skenario 5	56
Tabel 4.34 Confusion Matrix pada Fold 10 skenario 6	56
Tabel 4.35 Confusion Matrix pada Fold 10 skenario 7	57
Tabel 4.36 Confusion Matrix pada Fold 10 skenario 8	57
Tabel 4.37 Confusion Matrix pada Fold 10 skenario 9	58
Tabel 4.38 Confusion Matrix pada Fold 10 skenario 10	59
Tabel 4.39 rekapitulasi skenario dengan akurasi terbaik	59
Tabel 4.40 Data Validasi	61
Tabel 4.41 Komparasi hasil pengukuran Cross Fold dan data Validasi	62