

**IMPLEMENTASI METODE *MODIFIED K-NEAREST
NEIGHBOR* (MKNN) PADA PENENTUAN KEMINATAN
SEKOLAH MENGAH ATAS (SMA)
(STUDI KASUS : SMA NEGERI 1 SERIRIT)**

SKRIPSI

**Diajukan untuk memenuhi persyaratan
memperoleh gelar Sarjana Komputer**



**Disusun Oleh:
SITI HARDIYANTI
NIM. 105060800111051**

**KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS BRAWIJAYA
PROGRAM TEKNOLOGI INFORMASI DAN ILMU KOMPUTER
MALANG
2014**

LEMBAR PERSETUJUAN

IMPLEMENTASI METODE *MODIFIED K-NEAREST NEIGHBOR*

(MKNN) PADA PENENTUAN KEMINATAN

SEKOLAH MENGAH ATAS (SMA)

(STUDI KASUS : SMA NEGERI 1 SERIRIT)

SKRIPSI

Diajukan untuk memenuhi persyaratan memperoleh gelar Sarjana Komputer

Disusun Oleh:

SITI HARDIYANTI

NIM. 105060800111051

Menyetujui :

Pembimbing I

Pembimbing II

Muhammad Tanzil Furqon, S.Kom, M.CompSc
NIP. 19820930 200801 1 004

Indriati, S.T, M.Kom
NIK. 83101306120035

LEMBAR PENGESAHAN

IMPLEMENTASI METODE *MODIFIED K-NEAREST NEIGHBOR*

(MKNN) PADA PENENTUAN KEMINATAN

SEKOLAH MENGAH ATAS (SMA)

(STUDI KASUS : SMA NEGERI 1 SERIRIT)

SKRIPSI

Laboratorium Komputasi Cerdas dan Visualisasi

Diajukan untuk memenuhi persyaratan memperoleh gelar Sarjana Komputer

Disusun oleh:

SITI HARDIYANTI

NIM. 105060800111051

Skripsi ini telah diuji dan dinyatakan lulus tanggal 17 Oktober 2014

Penguji I

Penguji II

Candra Dewi, S.Kom, MSc
NIP. 19771114 200312 2 001

Edy Santoso, S.Si, M.Kom
NIP. 19740414 200312 1 004

Penguji III

Wijaya Kurniawan, S.T, M.T
NIK. 820125 16 1 1 0418

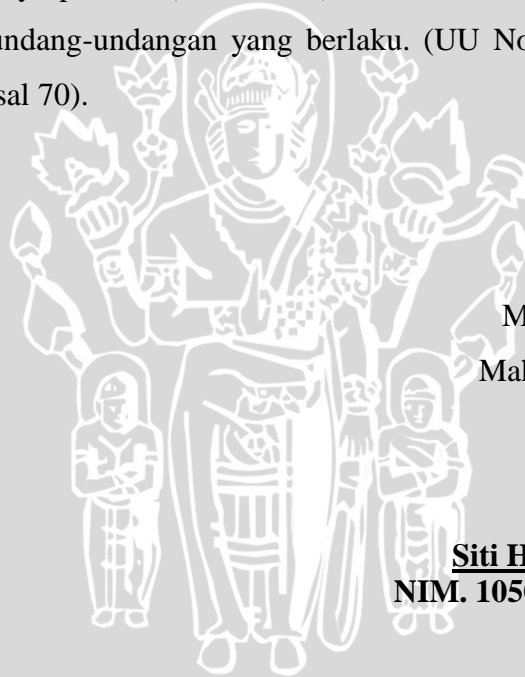
Mengetahui
Ketua Program Studi Informatika

Drs. Marji., M.T.
NIP. 19670801 199203 1 001

PERNYATAAN ORISINALITAS SKRIPSI

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam naskah SKRIPSI ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu perguruan tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip dalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata didalam naskah SKRIPSI ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia SKRIPSI ini digugurkan dan gelar akademik yang telah saya peroleh (SARJANA) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku. (UU No. 20 Tahun 2003, Pasal 25 ayat 2 dan Pasal 70).



Malang,
Mahasiswa,

Siti Hardiyanti
NIM. 105060800111051

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT, karena atas segala rahmat dan limpahan hidayah-Nya, penulis dapat menyelesaikan skripsi yang berjudul “Implementasi Metode *Modified K-Nearest Neighbor* (MKNN) Pada Penentuan Keminatan Sekolah Mengah Atas (SMA) (Studi Kasus : Sma Negeri 1 Seririt)”.

Dalam pelaksanaan dan penulisan skripsi ini penulis mendapatkan banyak bantuan dari berbagai pihak baik secara moral maupun material. Dalam kesempatan ini penulis ingin mengucapkan terima kasih yang sebesar – besarnya kepada :

1. Muhammad Tanzil Furqon, S.Kom, M.CompSc dan Indriati, S.T., M.Kom selaku dosen pembimbing selama pelaksanaan skripsi.
2. Ir. Sutrisno, M.T, Ir. Heru Nurwasito, M.Kom, Himawat Aryadita, S.T, M.Sc, dan Edy Santoso, S.Si, M.Kom selaku Ketua, Wakil Ketua 1, Wakil Ketua 2 dan Wakil Ketua 3 Program Teknologi Informasi dan Ilmu Komputer Universitas Brawijaya.
3. Drs. Marji, M.T dan Issa Arwani, S.Kom, M.Sc selaku Ketua dan Sekretaris Program Studi Teknik Informatika Universitas Brawijaya.
4. Ayahanda, Ibunda, dan seluruh keluarga atas segenap dukungan dan kasih sayang yang telah diberikan.
5. Seluruh Dosen Teknik Informatika Universitas Brawijaya atas kesediaan membagi ilmunya kepada penulis.
6. Seluruh Civitas Akademika Teknik Informatika Universitas Brawijaya yang telah banyak memberi bantuan dan dukungan selama penulis menempuh studi di Teknik Informatika Universitas Brawijaya dan selama penyelesaian skripsi ini.
7. Teman-teman angkatan 2010 dan Konsentrasi Cerdas dan Visualisasi dan seluruh pihak yang telah membantu kelancaran penulisan skripsi yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih banyak kekurangan baik format penulisan maupun isinya. Oleh karena itu, saran dan

kritik membangun senantiasa diharapkan guna perbaikan bagi penelitian selanjutnya. Semoga skripsi ini dapat memberikan manfaat bagi semua pihak.

Malang, 27 September 2014

Penulis



ABSTRAK

Berdasarkan kurikulum 2013, keminatan SMA tidak melewati tes dan hanya dilihat dari perolehan nilai mata pelajaran kelas IX SMP. Berdasarkan nilai matapelajaran tersebut, kemudian bisa disesuaikan kemana saja siswa yang bersangkutan bisa masuk sesuai dengan kemampuannya. Namun pada kenyataannya saat ini, banyak siswa memilih keminatan berdasarkan karena dorongan orang tua maupun mengikuti pilihan teman yang terkadang tidak sesuai dengan kemampuannya. Untuk mempermudah tahapan pengklasifikasian keminatan berdasarkan nilai rapor maka digunakan metode metode *Modified K-Nearest Neighbor* (MKNN). MKNN dilakukan dengan cara menghitung validitas data *training* sesuai dengan nilai k , menghitung jarak antara data uji dengan data training dan menghitung *weight voting*.

Hasil pengujian menunjukkan bahwa parameter nilai k sangat berpengaruh terhadap hasil klasifikasi dan akurasi yang dihasilkan. Rata-rata akurasi cenderung menurun seiring dengan penambahan nilai k sedangkan peningkatan jumlah data *training* turut disertai dengan peningkatan hasil akurasi. Untuk data *training* dengan sebaran kelas tidak seimbang mengalami penurunan nilai akurasi seiring dengan bertambahnya jumlah data. Hasil akurasi tertinggi didapat saat nilai $k=3$ yaitu 86.1% pada saat sebaran data kelas seimbang.

Kata Kunci: Kurikulum 2013, Keminatan SMA, *Modified K-Nearest Neighbor*.

ABSTRACT

Based on the curriculum of 2013, interest of the high school not pass the test and only seen from grades IX junior class lesson. Based on the value of the lesson, then can be adjusted anywhere can enter the student according to his ability. But in fact today, many students choose the interest thing based as well as urge parents to follow the choice that sometimes friends are not in accordance with his ability. To simplify the classification stages of specialization based on grades then used the method method *Modified K-Nearest Neighbor* (MKNN). MKNN done by calculating the validity of the training data according to the value of k , calculate the distance between the test data to the training data and calculate the weight of voting.

The test results showed that the value of the parameter k affects the classification results and the resulting accuracy. Average accuracy tends to decline with the addition of the value of k , while the increase in the amount of *training* data also accompanied by an increase in accuracy of the results, for the *training* data with imbalanced class accuracy values decreased with increasing amount of data. The results of the highest accuracy is obtained when the value of $k = 3$ is 86.1% upon the distribution of data by class.

Key : Curriculum of 2013, Interest of Senior High School, *Modified K-Nearest Neighbor*

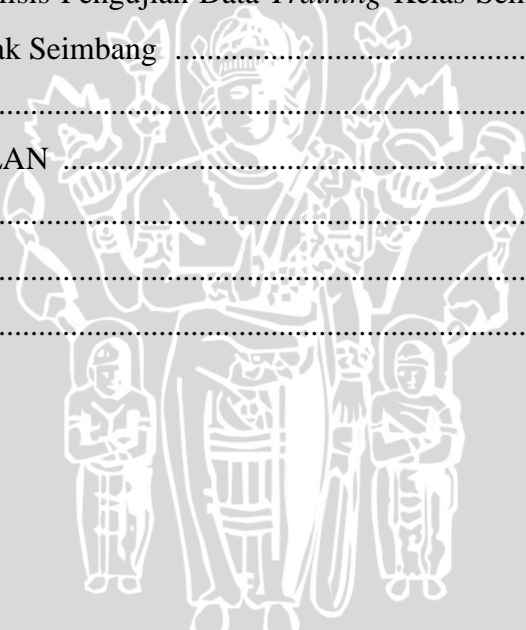
DAFTAR ISI

LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
PERNYATAAN ORISINALITAS SKRIPSI	iv
KATA PENGANTAR	v
ABSTRAK	vii
<i>ABSTRACT</i>	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiii
DAFTAR <i>SOURCE CODE</i>	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
1.1 LATAR BELAKANG	1
1.2 RUMUSAN MASALAH	2
1.3 BATASAN MASALAH	2
1.4 TUJUAN PENELITIAN	2
1.5 MANFAAT PENELITIAN	3
1.6 SISTEMATIKA PENULISAN	3
BAB II KAJIAN PUSTAKA DAN DASAR TEORI	5
2.1 KAJIAN PUSTAKA	5
2.2 KEMINATAN SMA	5
2.3 DATA	6
2.3.1 Definisi Data	6
2.4 <i>DATA MINING</i>	8
2.4.1 Definisi <i>Data Mining</i>	8
2.4.2 Proses <i>Data Mining</i>	9
2.4.3 Metode <i>Data Mining</i>	9
2.5 KLASIFIKASI	10
2.6 METODE KNN (K-NEAREST NEIGHBOR)	11



2.7 METODE <i>MODIFIED K-NEAREST NEIGHNOR</i> (MKNN)	13
2.8 AKURASI SISTEM	15
2.9 PENGENALAN PHP	15
2.10 MySQL	16
BAB III METODOLOGI PENELITIAN DAN PERANCANGAN	18
3.1 METODE PENELITIAN	18
3.1.1 Studi Literatur	19
3.1.2 Pengambilan Data	19
3.1.3 Perancangan Sistem	19
3.1.4 Implementasi Sistem	19
3.1.5 Pengujian Dan Analisi	19
3.1.6 Pengambilan Kesimpulan	19
3.2 ANALISI KEBUTUHAN	20
3.2.1 Deskripsi Umum	20
3.2.2 Deskripsi Data	20
3.3 PERANCANGAN PROSES	21
3.3.1 Proses Perhitungan Validitas	22
3.3.2 Menghitung Jarak (<i>Euclidean</i>)	24
3.3.3 Menghitung <i>Weight Voting</i>	24
3.4 PERHITUNGAN MANUAL	25
3.5 PERANCANGAN <i>DATABASE</i>	31
3.6 PERANCANGAN ANTARMUKA	33
3.7 PERANCANGAN UJI COBA	35
BAB IV IMPLEMENTASI	36
4.1 IMPLEMENTASI PROGRAM	36
4.1.1 Proses Menginputkan Data Dan Menyimpan Ke <i>Database</i> ..	36
4.1.2 Proses Perhitungan Validitas Data <i>Training</i>	40
4.1.3 Proses Perhitungan Jarak <i>Euclidean</i> Dan Normalisasi	41
4.1.4 Proses Perhitungan <i>Weight Voting</i>	42
4.2 IMPLEMENTASI ANTARMUKA	43
4.4.1 Tampilan Form Penginputan Data	43
4.4.2 Tampilan Form Hasil	44

BAB V PENGUJIAN DAN ANALISIS	45
5.1 PENGUJIAN	45
5.1.1 Pengujian Pengaruh Nilai k	45
5.1.2 Pengujian Pada Jumlah Data Testing Tetap Dengan Jumlah Data <i>Training</i> Berbeda	46
5.1.3 Pengujian Data <i>Training</i> Kelas Seimbang Dan Kelas Tidak Seimbang	47
5.2 ANALISIS	48
5.2.1 Analisis Pengaruh Nilai k	48
5.2.2 Analisis Pengujian Pada Jumlah Data Testing Tetap Dengan Jumlah Data <i>Training</i> Berbeda	48
5.2.3 Analisis Pengujian Data <i>Training</i> Kelas Seimbang Dan Kelas Tidak Seimbang	49
BAB VI PENUTUP	51
6.1 KESIMPULAN	51
6.2 SARAN	51
DAFTAR PUSTAKA	DP-1
LAMPIRAN	L-1



DAFTAR GAMBAR

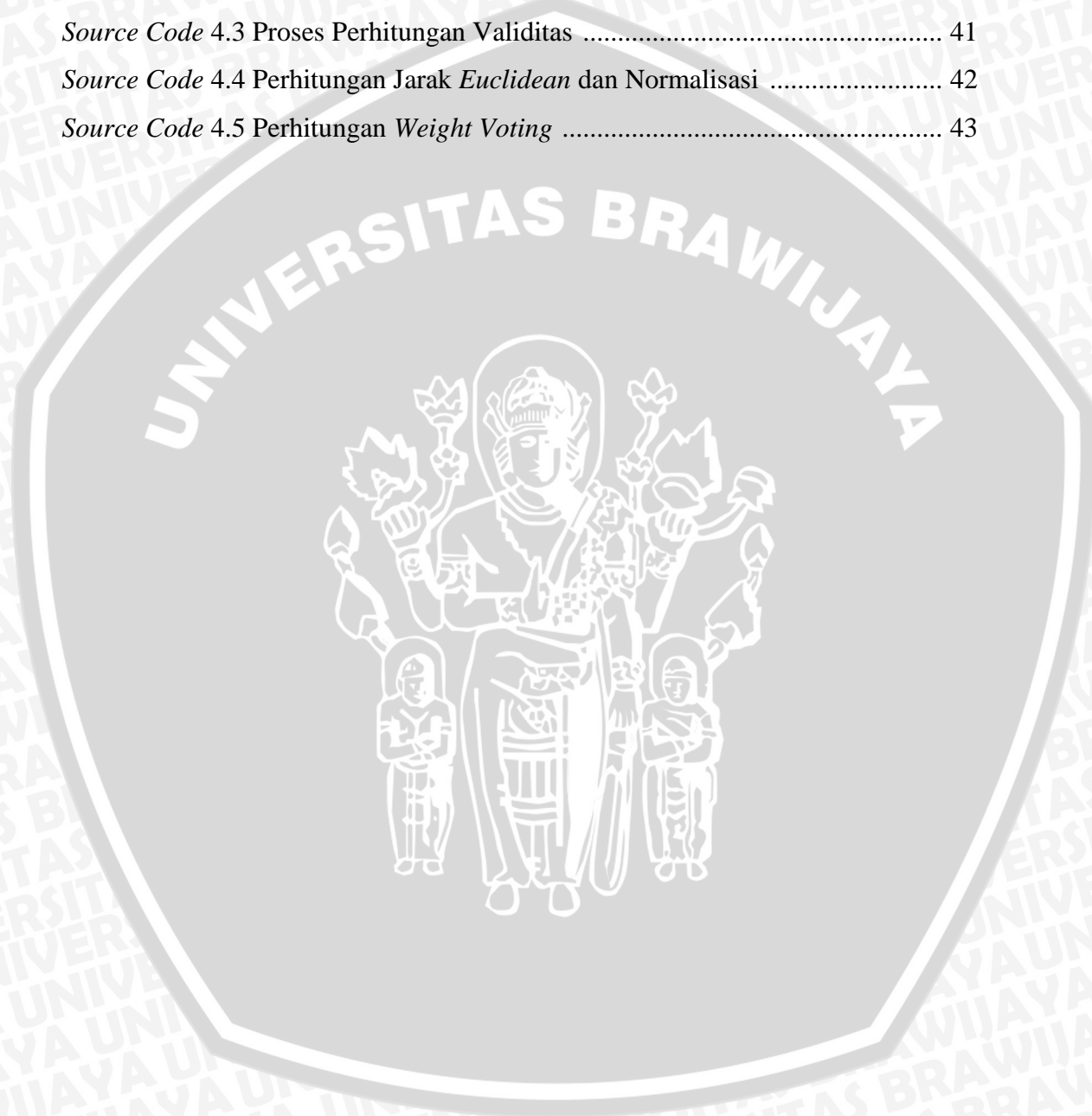
Gambar 2.1 Proses <i>Knowledge Discovery in Database</i> (KDD)	6
Gambar 2.2 <i>Flowchart</i> dari Metode KNN	13
Gambar 3.1 Alur Penelitian	18
Gambar 3.2 <i>Flowchart</i> Sistem	20
Gambar 3.3 <i>Flowchart</i> Proses MKNN	21
Gambar 3.4 <i>Flowchart</i> Proses Perhitungan Jarak <i>Euclidean Data Training</i>	22
Gambar 3.5 <i>Flowchart</i> Proses Perhitungan Validitas	23
Gambar 3.6 <i>Flowchart</i> Proses Perhitungan <i>Euclidean</i>	24
Gambar 3.7 <i>Flowchart</i> Proses Perhitungan <i>Weight Voting</i>	25
Gambar 3.8 Rancangan <i>Database</i>	31
Gambar 3.9 Rancangan Antarmuka Home	34
Gambar 3.10 Rancangan Proses	34
Gambar 4.1 Tampilan Form Penginputan Data	44
Gambar 4.2 Tampilan Form Hasil	44
Gambar 5.1 Grafik Pengaruh Nilai k Terhadap Akurasi	48
Gambar 5.2 Grafik Penambahan Jumlah Data <i>Training</i> Terhadap Akurasi	49
Gambar 5.3 Grafik Pengaruh Data Training Kelas Seimbang dan Kelas Tidak Seimbang Terhadap Akurasi	50

DAFTAR TABEL

Tabel 3.1 Dataset Perhitungan Manual	26
Tabel 3.2 Data Testing	27
Tabel 3.3 Data <i>Training</i>	27
Tabel 3.4 Perhitungan <i>Euclidean Data Training</i>	28
Tabel 3.5 Perhitungan Validitas	28
Tabel 3.6 Perhitungan <i>Euclidean Data Testing 1</i>	30
Tabel 3.7 Hasil Perhitungan <i>Weight Voting Data Testing 1</i>	30
Tabel 3.8 Hasil Perhitungan 3 Data Testing	31
Tabel 3.9 Tabel data_siswa	32
Tabel 3.10 Tabel validitas	32
Tabel 3.11 Tabel jarak_data	33
Tabel 3.12 Tabel jarak_data_normalisasi	33
Tabel 3.13 Tabel hasil	33
Tabel 3.14 Rancangan Tabel Pengujian Akurasi	35
Tabel 5.1 Hasil Uji Coba Pengaruh Nilai k Dataset Menggunakan Rumus	45
Tabel 5.2 Hasil Uji Coba Pengaruh Nilai k Dataset Asli	46
Tabel 5.3 Hasil Pengujian dengan Jumlah Data Testing Tetap dengan Jumlah Data <i>Training</i> Berbeda	47
Tabel 5.4 Hasil Pengujian Kelas Seimbang dan Tidak Seimbang	47

DAFTAR SOURCE CODE

<i>Source Code 4.1</i> Proses Menginputkan Data	38
<i>Source Code 4.2</i> Proses Menyimpan Data ke <i>Database</i>	39
<i>Source Code 4.3</i> Proses Perhitungan Validitas	41
<i>Source Code 4.4</i> Perhitungan Jarak <i>Euclidean</i> dan Normalisasi	42
<i>Source Code 4.5</i> Perhitungan <i>Weight Voting</i>	43



DAFTAR LAMPIRAN

Lampiran 1 Dataset Asli SMA Negeri 1 Seririt	L-1
Lampiran 2 Dataset Keminatan Berdasarkan Rumus	L-9
Lampiran 3 Hasil Uji Coba Pengaruh Nilai k Kelas Seimbang	L-17
Lampiran 4 Hasil Uji Coba Pengaruh Nilai k Kelas Tidak Seimbang	L-29
Lampiran 5 Hasil Uji Coba Pengaruh Nilai k Dataset Asli	L-21

