

**OPTIMASI METODE NAIVE BAYES DENGAN ALGORITMA
GENETIKA UNTUK KLASIFIKASI KETEPATAN WAKTU
STUDI MAHASISWA**

*Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-I Pada
Jurusan Teknik Informatika*



Oleh:

Elsya Krismi Afindri
NIM : 09021281419061

**Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2018**

LEMBAR PENGESAHAN TUGAS AKHIR

OPTIMASI METODE NAIVE BAYES DENGAN ALGORITMA
GENETIKA UNTUK KLASIFIKASI KETEPATAN WAKTU
STUDI MAHASISWA

Oleh :

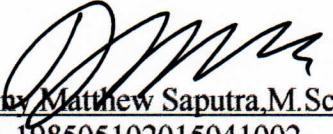
ELSYA KRISMI AFINDRI
NIM : 09021281419061

Palembang, 9 November 2018

Pembimbing I,


Yoppy Sazaki, M.T
NIP. 197406062012101201

Pembimbing II,


Danny Matthew Saputra, M.Sc
NIP. 198505102015041002

Mengetahui,
Ketua Jurusan Teknik Informatika



TANDA LULUS SIDANG SKRIPSI

Pada hari Kamis, tanggal 25 Oktober 2018 telah dilaksanakan ujian sidang skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Elsy Krismi Afindri

NIM : 09021281419061

Judul : Optimasi Metode Naive Bayes dengan Algoritma Genetika untuk Klasifikasi Ketepatan Waktu Studi Mahasiswa

1. Pembimbing I

Yoppy Sazaki, M.T
NIP. 197406062012101201



2. Pembimbing II

Danny Matthew Saputra, M.Sc
NIP. 198505102015041002



3. Penguji I

Dian Palupi Rini, M.Kom., Ph.D
NIP. 197802232006042002



4. Penguji II

Ir. M. Ihsan Jambak, M.Sc
NIP. 196804052013081201



Mengetahui,
Ketua Jurusan Teknik Informatika



Rifkie Primartha, M.T
NIP. 197706012009121004

HALAMAN PERNYATAAN BEBAS PLAGIAT

Yang bertanda tangan di bawah ini :

Nama : Elsyia Krismi Afindri
NIM : 090212181419061
Program Studi : Teknik Informatika Reguler
Judul Skripsi : Optimasi Metode Naive Bayes dengan Algoritma Genetika untuk Klasifikasi Ketepatan Waktu Studi Mahasiswa

Hasil Pengecekan Software *iThenticate/Turnitin* : 16 %

Menyatakan bahwa Laporan Projek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan projek ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, 9 November 2018



Elsya Krismi Afindri
NIM. 090212181419061

*There is no elevator to success, you
have to take the stairs.*

Kupersembahkan karya tulis ini kepada :

- ❖ Orang tuaku
- ❖ Adik-adik kesayanganku
- ❖ Teman dekatku
- ❖ Rekan-rekan seperjuanganku
- ❖ Almamater kebanggaanku

OPTIMIZATION OF NAIVE BAYES METHOD WITH GENETIC
ALGORITHMS TO CLASSIFY THE TIMELINESS OF STUDENT STUDY

by:

Elsya Krismi Afindri
09021281419061

ABSTRACT

Naive Bayes is a very effective and widely used classification method, in this method, there are independences of features that assume all features are important and need to be calculated individually to determine the results of the classification, this affects the accuracy of classification. This research will optimize the classification by Naive Bayes using Genetic Algorithm to do feature selection. Feature selection used to obtain important features to be calculated for determine the data classification and increase the accuracy of the classification. The data used are the data from Informatics students of Sriwijaya University from class of 2011 until 2013. The classification that done here is the classification of timeliness of student study. The result of this research shows the average accuracy of classification is 83%, which has an average increase by 8% from the accuracy of classification without optimization, and the maximum accuracy of classification that can be reached by optimization is 84.12% with 27 selected features. The Genetic Algorithm gives a better effect on the results of Naïve Bayes Classification by doing feature selection that increase the accuracy.

Keywords: Classification, Feature Selection, Naive Bayes, Genetic Algorithms

OPTIMASI METODE NAIVE BAYES DENGAN ALGORITMA GENETIKA
UNTUK KLASIFIKASI KETEPATAN WAKTU STUDI MAHASISWA

Oleh:

Elsya Krismi Afindri
09021281419061

ABSTRAK

Naive Bayes merupakan metode klasifikasi yang sangat efektif dan banyak digunakan, di dalam metode ini terdapat independensi atribut yang menganggap semua atribut penting dan perlu diperhitungkan secara individual untuk menentukan hasil klasifikasi suatu data, hal ini berpengaruh pada nilai akurasi yang dihasilkan. Penelitian ini akan melakukan optimasi terhadap hasil klasifikasi metode Naive Bayes dengan seleksi fitur menggunakan Algoritma Genetika. Seleksi fitur dilakukan untuk mendapatkan fitur-fitur penting yang akan dihitung probabilitasnya untuk menentukan klasifikasi data dan dapat meningkatkan nilai akurasi dari hasil klasifikasi. Data yang digunakan merupakan data mahasiswa Teknik Informatika Reguler Universitas Sriwijaya Angkatan 2011-2013. Klasifikasi yang dilakukan adalah klasifikasi ketepatan waktu studi mahasiswa. Penelitian ini menghasilkan nilai akurasi rata-rata sebesar 83%, yang mana mengalami peningkatan rata-rata sebesar 8% dari nilai akurasi sebelum optimasi, dan nilai akurasi maksimal yang bisa didapatkan ketika optimasi dilakukan mencapai 84,12% dengan 27 atribut terpilih. Algoritma Genetika memberikan pengaruh yang lebih baik terhadap hasil klasifikasi Naïve Bayes dengan melakukan seleksi fitur yang meningkatkan nilai akurasi.

Kata Kunci: Klasifikasi, Seleksi Fitur, Naive Bayes, Algoritma Genetika

**OPTIMIZATION OF NAIVE BAYES METHOD WITH GENETIC
ALGORITHMS TO CLASSIFY THE TIMELINESS OF STUDENT STUDY**

by:

Elsya Krismi Afindri
09021281419061

ABSTRACT

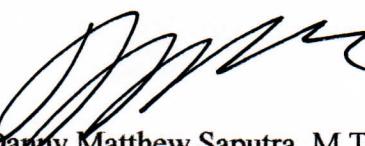
Naive Bayes is a very effective and widely used classification method, in this method, there are independences of features that assume all features are important and need to be calculated individually to determine the results of the classification, this affects the accuracy of classification. This research will optimize the classification by Naive Bayes using Genetic Algorithm to do feature selection. Feature selection used to obtain important features to be calculated for determine the data classification and increase the accuracy of the classification. The data used are the data from Informatics students of Sriwijaya University from class of 2011 until 2013. The classification that done here is the classification of timeliness of student study. The result of this research shows the average accuracy of classification is 83%, which has an average increase by 8% from the accuracy of classification without optimization, and the maximum accuracy of classification that can be reached by optimization is 84.12% with 27 selected features. The Genetic Algorithm gives a better effect on the results of Naïve Bayes Classification by doing feature selection that increase the accuracy.

Keywords: Classification, Feature Selection, Naive Bayes, Genetic Algorithms

Pembimbing I,


Yoppy Sasaki, M.T
NIP. 196804052013081201

Indralaya, November 2018
Pembimbing II,


Danny Matthew Saputra, M.T
NIP. 198505102015041002

Mengetahui,
Ketua Jurusan Teknik Informatika



OPTIMASI METODE NAIVE BAYES DENGAN ALGORITMA GENETIKA
UNTUK KLASIFIKASI KETEPATAN WAKTU STUDI MAHASISWA

Oleh:

Elsya Krismi Afindri
09021281419061

ABSTRAK

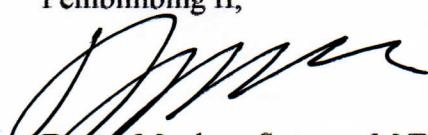
Naive Bayes merupakan metode klasifikasi yang sangat efektif dan banyak digunakan, di dalam metode ini terdapat independensi atribut yang menganggap semua atribut penting dan perlu diperhitungkan secara individual untuk menentukan hasil klasifikasi suatu data, hal ini berpengaruh pada nilai akurasi yang dihasilkan. Penelitian ini akan melakukan optimasi terhadap hasil klasifikasi metode Naive Bayes dengan seleksi fitur menggunakan Algoritma Genetika. Seleksi fitur dilakukan untuk mendapatkan fitur-fitur penting yang akan dihitung probabilitasnya untuk menentukan klasifikasi data dan dapat meningkatkan nilai akurasi dari hasil klasifikasi. Data yang digunakan merupakan data mahasiswa Teknik Informatika Reguler Universitas Sriwijaya Angkatan 2011-2013. Klasifikasi yang dilakukan adalah klasifikasi ketepatan waktu studi mahasiswa. Penelitian ini menghasilkan nilai akurasi rata-rata sebesar 83%, yang mengalami peningkatan rata-rata sebesar 8% dari nilai akurasi sebelum optimasi, dan nilai akurasi maksimal yang bisa didapatkan ketika optimasi dilakukan mencapai 84,12% dengan 27 atribut terpilih. Algoritma Genetika memberikan pengaruh yang lebih baik terhadap hasil klasifikasi Naïve Bayes dengan melakukan seleksi fitur yang meningkatkan nilai akurasi.

Kata Kunci: Klasifikasi, Seleksi Fitur, Naive Bayes, Algoritma Genetika

Pembimbing I,


Yoppy Sazaki, M.T
NIP. 196804052013081201

Indralaya, November 2018
Pembimbing II,


Danny Matthew Saputra, M.T
NIP. 198505102015041002



KATA PENGANTAR



Puji syukur kepada Allah SWT atas berkat dan rahmat-Nya yang telah diberikan kepada Penulis sehingga dapat menyelesaikan Tugas Akhir ini dengan baik. Tugas akhir ini disusun untuk memenuhi salah satu syarat guna menyelesaikan pendidikan program Strata-1 pada Fakultas Ilmu Komputer Program Studi Teknik Informatika di Universitas Sriwijaya.

Dalam menyelesaikan Tugas Akhir ini banyak pihak yang telah memberikan bantuan dan dukungan baik secara langsung maupun secara tidak langsung. Penulis ingin menyampaikan rasa terima kasih kepada:

1. Bapak Jaidan Jauhari, S.Pd., M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
2. Bapak Rifkie Primartha, S.T., M.T. selaku Ketua Jurusan Teknik Informatika yang selalu memberikan pembinaan dalam proses perkuliahan dan penggerjaan Tugas Akhir.
3. Papa, mama, dan adik-adik saya yang selalu memberikan dukungan baik moril maupun materil, serta keluarga besar yang selalu mendoakan.
4. Bapak Yoppy Sazaki, M.T selaku dosen pembimbing I yang telah membimbing, mengarahkan, dan memberikan motivasi dalam proses perkuliahan dan penggerjaan Tugas Akhir.
5. Bapak Danny Matthew Saputra, M.Sc selaku dosen pembimbing II yang telah membimbing, mengarahkan, dan selalu mendukung saya dengan topik yang saya ambil untuk Tugas Akhir.
6. Ibu Dian Palupi Rini, M.Kom., Ph.D selaku dosen penguji I yang telah memberikan masukan dan dorongan dalam proses penggerjaan Tugas Akhir.

7. Bapak Ir. M. Ihsan Jambak, M.Sc selaku dosen penguji II yang telah memberikan masukan dan dorongan dalam proses penggerjaan Tugas Akhir.
8. Seluruh dosen Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
9. Mbak Winda dan Mbak Wiwin yang telah membantu dalam kelancaran proses administrasi dan akademik selama masa perkuliahan, serta seluruh staf administrasi yang telah membantu dalam pengurusan surat dan berkas lainnya.
10. Abu yang 24/7 selalu ada dan siap sedia untuk membantu, memberikan dukungan, serta meyakinkan diri saya untuk dapat menyelesaikan Tugas Akhir ini.
11. Fahri yang selalu mau membantu saat diperlukan. Inayah dan Azhary yang mau membagikan ilmunya untuk membantu penggerjaan Tugas Akhir ini, serta teman-teman IF Reguler 2014 yang sudah bersama-sama berjuang dalam menempuh ilmu sejak awal perkuliahan.
12. Teman-teman anggota BEMF 2015, BPH BEMF 2016, dan inti BEMF 2017 yang memberikan saya pengalaman berorganisasi selama perkuliahan.
13. Diri saya sendiri yang telah berhasil untuk bertahan dan dapat menyelesaikan Tugas Akhir ini.

Penulis menyadari dalam penyusunan Tugas Akhir ini masih terdapat banyak kekurangan disebabkan keterbatasan pengetahuan dan pengalaman, oleh karena itu kritik dan saran yang membangun sangat diharapkan untuk kemajuan penelitian selanjutnya. Akhir kata semoga Tugas Akhir ini dapat berguna dan bermanfaat bagi kita semua.

Indralaya, November 2018

Elsya Krismi Afindri

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERSETUJUAN KOMISI PENGUJI	iii
HALAMAN PERNYATAAN BEBAS PLAGIAT	iv
HALAMAN MOTTO DAN PERSEMBAHAN.....	v
ABSTRACT	vi
ABSTRAK	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xvi
DAFTAR GAMBAR	xix
DAFTAR LAMPIRAN.....	xxii

BAB I PENDAHULUAN

1.1 Pendahuluan.....	I – 1
1.2 Latar Belakang Masalah	I – 1
1.3 Rumusan Masalah.....	I – 3
1.4 Tujuan Penelitian	I – 3
1.5 Manfaat Penelitian	I – 4
1.6 Batasan Masalah	I – 4
1.7 Sistematika Penulisan	I – 5
1.8 Kesimpulan	I – 6

BAB II KAJIAN LITERATUR

2.1 Pendahuluan.....	II – 1
2.2 Klasifikasi	II – 1
2.3 Naive Bayes	II – 4
2.4 Seleksi Fitur	II – 7
2.5 Algoritma Genetika	II – 8
2.6 Penelitian Lain yang Relevan	II – 10
2.7 Kesimpulan	II – 14

BAB III METODOLOGI PENELITIAN

3.1 Pendahuluan.....	III – 1
3.2 Unit Penelitian	III – 1
3.3 Pengumpulan Data	III – 1
3.4 Tahapan Penelitian.....	III – 2
3.4.1 Menetapkan Kerangka Kerja	III – 2
a. Praproses.....	III – 2
b. Klasifikasi Naive Bayes	III – 3
c. Algoritma Genetika	III – 4
d. <i>K-Fold Cross Validation</i>	III – 7
d. <i>Confusion Matrix</i>	III – 7
3.4.2 Menetapkan Kriteria Pengujian	III – 8
3.4.3 Menetapkan Format Data Pengujian.....	III – 8

3.4.4 Menentukan Alat yang Digunakan dalam Pelaksanaan Penelitian	III – 10
3.4.5 Melakukan Pengujian Penelitian.....	III – 10
3.4.6 Melakukan Analisa Hasil Pengujian dan Membuat Kesimpulan.....	III – 12
3.5 Metode Pengembangan Perangkat Lunak.....	III – 14
3.5.1 Rational Unified Process.....	III – 14
3.5.2 Fase Insepsi	III – 15
3.5.3 Fase Elaborasi	III – 15
3.5.4 Fase Konstruksi.....	III – 16
3.5.5 Fase Transisi	III – 16
3.6 Manajemen Proyek Penelitian	III – 17

BAB IV PENGEMBANGAN PERANGKAT LUNAK

4.1 Pendahuluan.....	IV – 1
4.2 Fase Insepsi	IV – 1
4.2.1 Pemodelan Bisnis.....	IV – 1
4.2.2 Kebutuhan Sistem	IV – 2
4.2.2.1 Fitur Praproses	IV – 3
4.2.2.2 Fitur Seleksi Fitur dengan Algoritma Genetika	IV – 3
4.2.2.3 Fitur Klasifikasi	IV – 3
4.2.3 Analisis dan Desain.....	IV – 4
4.2.3.1 Analisis Kebutuhan Perangkat Lunak.....	IV – 5

4.2.3.2 Analisis Data.....	IV – 6
4.2.3.3 Analisis Praproses	IV – 6
a. Data Transformation.....	IV – 8
b. Data Discretization	IV – 9
4.2.3.4 Analisis Metode Naïve Bayes	IV – 10
4.2.3.5 Analisis Metode Algoritma Genetika	IV – 11
4.2.3.6 Desain Perangkat Lunak	IV – 15
1. Model Use Case.....	IV – 15
2. Diagram Aktivitas	IV – 22
4.3 Fase Elaborasi	IV – 25
4.3.1 Pemodelan Bisnis.....	IV – 25
4.3.1.1 Perancangan Data.....	IV – 25
4.3.1.2 Perancangan Antar Muka.....	IV – 26
4.3.2 Kebutuhan Sistem	IV – 28
4.3.3 Diagram Sequence	IV – 29
4.4 Fase Konstruksi.....	IV – 33
4.4.1 Kebutuhan Sistem	IV – 33
4.4.2 Diagram Kelas	IV – 33
4.4.3 Implementasi.....	IV – 35
4.4.3.1 Implementasi Kelas.....	IV – 35
4.4.3.2 Implementasi Antarmuka.....	IV – 36
4.5 Fase Transisi	IV – 42
4.5.1 Pemodelan Bisnis.....	IV – 42

4.5.2 Kebutuhan Sistem	IV – 42
4.5.3 Rencana Pengujian.....	IV – 43
4.5.3.1 Rencana Pengujian Use Case Melakukan Praproses Data..	IV – 43
4.5.3.2 Rencana Pengujian Use Case Melakukan Seleksi fitur dengan Algoritma Genetika	IV – 43
4.5.3.3 Rencana Pengujian Use Case Melakukan Klasifikasi dengan Naive Bayes	IV – 43
4.5.4 Implementasi.....	IV – 46
4.5.4.1 Pengujian Use Case Melakukan Praproses Data	IV – 46
4.5.4.2 Rencana Pengujian Use Case Melakukan Seleksi fitur dengan Algoritma Genetika	IV – 47
4.5.4.3 Rencana Pengujian Use Case Melakukan Klasifikasi dengan Naive Bayes	IV – 49

BAB V ANALISIS PENELITIAN

5.1 Pendahuluan.....	V – 1
5.2 Hasil Percobaan Penelitian	V – 1
5.2.1 Hasil Klasifikasi Naive Bayes	V – 3
5.2.2 Hasil Klasifikasi Naive Bayes dan Algoritma Genetika.....	V – 5
5.3 Confussion Matrix	V – 8
5.4 Atribut Terpilih	V – 10
5.5 Nilai Akurasi, Precision, dan Recall	V – 12
5.6 Analisis Penelitian	V – 17

5.6.1 Uji T Dua Sampel	V – 22
5.7 Perbandingan Hasil Penelitian	V – 25
5.8 Kesimpulan	V – 28

BAB VI KESIMPULAN DAN SARAN

6.1 Pendahuluan.....	VI – 1
6.2 Kesimpulan	VI – 1
6.3 Saran	VI – 2

DAFTAR PUSTAKA	xxiii
----------------------	-------

DAFTAR TABEL

	Halaman
Tabel III-1. Rancangan Tabel Hasil Pengujian Klasifikasi dengan Naive Bayes.....	III – 9
Tabel III-2. Rancangan Tabel Hasil Pengujian Algoritma Genetika pada metode Naive Bayes untuk Klasifikasi Ketepatan Waktu Studi Mahasiswa	III – 9
Tabel III-3. Rancangan Tabel Analisa Hasil Klasifikasi dengan Naive Bayes	III – 12
Tabel III-3. Rancangan Tabel Perbandingan Hasil Pengujian Klasifikasi metode Naive Bayes dengan dan tanpa menggunakan Algoritma Genetika	III – 12
Tabel IV-1. Kebutuhan Fungsional.....	IV – 4
Tabel IV-2. Kebutuhan Non Fungsional	IV – 4
Tabel IV-3. Contoh Data Mahasiswa sebelum Praproses	IV – 7
Tabel IV-4. Contoh Data Mahasiswa setelah Data Transformation	IV – 8
Tabel IV-5. Contoh Data Mahasiswa setelah Data Discretization.....	IV – 9
Tabel IV-6. Hasil Klasifikasi Contoh Data Mahasiswa	IV – 10
Tabel IV-7. Contoh Perhitungan Algoritma Genetika	IV – 11
Tabel IV-8. Definisi Aktor <i>Use Case</i>	IV – 15
Tabel IV-9. Definisi <i>Use Case</i>	IV – 17
Tabel IV-10. Skenario <i>Use Case</i> Melakukan Praproses Data.....	IV – 18

Tabel IV-11. Skenario <i>Use Case</i> Melakukan Seleksi Fitur dengan Algoritma Genetika.....	IV – 19
Tabel IV-12. Skenario <i>Use Case</i> Melakukan Klasifikas dengan Naive Bayes.....	IV – 21
Tabel IV-13. Implementasi Kelas	IV – 35
Tabel IV-14. Rencana Pengujian <i>Use Case</i> Melakukan Praproses Data.....	IV – 43
Tabel IV-15. Rencana Pengujian <i>Use Case</i> Melakukan Seleksi Fitur dengan Algoritma Genetika	IV – 43
Tabel IV-16. Rencana Pengujian <i>Use Case</i> Melakukan Klasifikasi dengan Naive Bayes.....	IV – 45
Tabel IV-17. Pengujian <i>Use Case</i> Melakukan Praproses Data.....	IV – 46
Tabel IV-18. Pengujian <i>Use Case</i> Melakukan Seleksi Fitur dengan Algoritma Genetika.....	IV – 47
Tabel IV-19. Pengujian <i>Use Case</i> Melakukan Klasifikasi dengan Naive Bayes	IV – 49
Tabel V-1. Hasil Klasifikasi Naïve Bayes untuk Fold ke-1	V – 4
Tabel V-2. Hasil Klasifikasi Naïve Bayes dan Algoritma Genetika.....	V – 5
Tabel V-3. Hasil Klasifikasi Naïve Bayes dan Algoritma Genetika dengan parameter optimal	V – 7
Tabel V-4. Confussion Matrix	V – 8
Tabel V-5. Confussion Matrix Hasil Klasifikasi Naïve Bayes	V – 8
Tabel V-6. Confussion Matrix Hasil Klasifikasi Naïve Bayes dan Algoritma Genetika sebagai seleksi fitur	V – 9

Tabel V-7. Confussion Matrix Hasil Klasifikasi Naïve Bayes dan Algoritma Genetika sebagai seleksi fitur dengan parameter optimal.....	V – 10
Tabel V-8. Perbandingan Jumlah Fitur dan Nilai Akurasi untuk Seleksi Fitur dengan Algoritma Genetika (Parameter Berbeda).....	V – 11
Tabel V-9. Perbandingan Jumlah Fitur dan Nilai Akurasi untuk Seleksi Fitur dengan Algoritma Genetika (Parameter Optimal).....	V – 12
Tabel V-10. Evaluasi Performa Klasifikasi Naïve Bayes tanpa seleksi fitur..	V – 13
Tabel V-11. Nilai Akurasi, Precision, dan Recall Klasifikasi Naïve Bayes dengan seleksi fitur menggunakan Algoritma Genetika.....	V – 14
Tabel V-12. Nilai Akurasi, Precision, dan Recall Klasifikasi Naïve Bayes dengan seleksi fitur menggunakan Algoritma Genetika dengan parameter optimal.....	V – 16
Tabel V-13. Tabel Hasil Uji Normalitas Seluruh Variabel Dependen terhadap Faktor Pembanding	V – 18
Tabel V-14. Tabel Hasil Uji Homogenitas Seluruh Variabel Dependen terhadap Faktor Pembanding	V – 20
Tabel V-15. Tabel Hasil Uji T untuk Dua Sampel Homogenitas Seluruh Variabel Dependen	V – 23

DAFTAR GAMBAR

	Halaman
Gambar II-1. Contoh Proses Pembelajaran	II – 3
Gambar II-2. Contoh Proses Klasifikasi	II – 3
Gambar III-1. Diagram Tahapan Pengujian Penelitian.....	III – 10
Gambar III-2. Diagram Tahapan Klasifikasi dengan Naive Bayes.....	III – 11
Gambar III-3. Diagram Tahapan Seleksi Fitur dengan Algoritma Genetika .	III – 11
Gambar III-4. Arsitektur RUP.....	III – 14
Gambar III-5. Penjadwalan untuk Tahap Menentukan Ruang Lingkup dan Unit Penelitian	III – 18
Gambar III-6. Penjadwalan untuk Tahap Menentukan Dasar Teori yang Berkaitan dengan Penelitian dan Menentukan Kriteria Pengujian.....	III – 18
Gambar III-7. Penjadwalan untuk Tahap Menentukan Alat yang Digunakan untuk Pelaksanaan Penelitian Fase Insepsi.....	III – 19
Gambar III-8. Penjadwalan untuk Tahap Menentukan Alat yang Digunakan untuk Pelaksanaan Penelitian Fase Elaborasi	III – 20
Gambar III-9. Penjadwalan untuk Tahap Menentukan Alat yang Digunakan untuk Pelaksanaan Penelitian Fase Konstruksi	III – 20
Gambar III-10. Penjadwalan untuk Tahap Menentukan Alat yang Digunakan untuk Pelaksanaan Penelitian Fase Transisi	III – 21

Gambar III-11. Penjadwalan untuk Tahap Melakukan Pengujian Penelitian, Analisa Hasil Pengujian Penelitian dan Membuat Kesimpulan	III – 21
Gambar IV-1. Predikat Kelulusan Universitas Sriwijaya	IV – 7
Gambar IV-2. Diagram Use Case	IV – 16
Gambar IV-3. Diagram Aktivitas Use Case Melakukan Praproses Data.....	IV – 22
Gambar IV-4. Diagram Aktivitas Use Case Melakukan Seleksi Fitur dengan Algoritma Genetika.....	IV – 23
Gambar IV-5. Diagram Aktivitas Use Case Melakukan Klasifikasi dengan Naive Bayes	IV – 24
Gambar IV-6. Rancangan Antarmuka Menu Utama.....	IV – 26
Gambar IV-7. Rancangan Antarmuka Menu Pilihan Klasifikasi.....	IV – 27
Gambar IV-8. Rancangan Antarmuka Menu Hasil Klasifikasi	IV – 28
Gambar IV-9. Diagram Sequence Melakukan Praproses Data.....	IV – 30
Gambar IV-10. Diagram Sequence Melakukan Seleksi Fitur dengan Algoritma Genetika.....	IV – 31
Gambar IV-11. Diagram Sequence Melakukan Klasifikasi dengan Naive Bayes	IV – 32
Gambar IV-12. Diagram Kelas Perangkat Lunak	IV – 34
Gambar IV-13. Antarmuka Halaman Awal Perangkat Lunak	IV – 37
Gambar IV-14. Antarmuka Halaman Utama	IV – 38
Gambar IV-15. Antarmuka Halaman Pilihan Klasifikasi	IV – 39

Gambar IV-16. Antarmuka Halaman Hasil Klasifikasi dengan Naïve Bayes	IV – 40
Gambar IV-17. Antarmuka Halaman Masukan Parameter Algoritma Genetika	IV – 40
Gambar IV-18. Antarmuka Halaman Hasil Klasifikasi Naïve Bayes dengan Seleksi Fitur Menggunakan Algoritma Genetika ...	IV – 41
Gambar V-1. Perbandingan Hasil Penelitian	V – 26

DAFTAR LAMPIRAN

Lampiran I. Penjadwalan Penelitian dalam Bentuk Work Breakdown Structure
(WBS)

Lampiran II. Contoh Perhitungan Naive Bayes

Lampiran III. Hasil Perhitungan Klasifikasi Naive Bayes untuk Seluruh Data Uji

Lampiran IV. Detail Atribut Terpilih untuk Seleksi Fitur dengan Algoritma
Genetika

Lampiran V. Kode Program PHP

Lampiran VI. Hasil Perhitungan Klasifikasi Naive Bayes dan Algoritma Genetika
untuk Data Uji pada Setiap Percobaan yang Memiliki Perbedaan
Hasil Klasifikasi

BAB I

PENDAHULUAN

1.1 Pendahuluan

Dalam bab ini membahas tentang latar belakang masalah, rumusan masalah, tujuan dan manfaat penelitian, serta batasan masalah. Bab ini akan memberikan uraian secara umum tentang keseluruhan penelitian.

Pendahuluan dimulai dengan penjelasan tujuan data mining, metode data mining dalam proses mendapatkan pengetahuan yaitu klasifikasi dengan Naive Bayes, serta penelitian yang berkaitan dengan penggabungan metode klasifikasi Naive Bayes dengan seleksi fitur menggunakan algoritma genetika.

1.2 Latar Belakang Masalah

Data Mining adalah proses mengetahui pola tertentu dan informasi dari data yang berjumlah sangat banyak dari sumber penyimpanan seperti database, data warehouse, internet, serta penyimpanan informasi lainnya, atau juga data yang sudah terhubung dengan sebuah sistem secara dinamik (Han, Pei, and Kamber, 2011). Data Mining dilakukan untuk mendapatkan informasi yang berguna dari sekumpulan data yang sangat banyak agar data yang ada dapat dipelajari dan memberikan ilmu/*knowledge* baru yang bisa dimanfaatkan. Salah satu cara yang digunakan untuk melakukan data mining adalah klasifikasi. Klasifikasi adalah sebuah rangkaian analisis data yang menghasilkan model untuk mendeskripsikan sebuah kelas data (Han, Pei, and Kamber, 2011).

Naive Bayes merupakan salah satu metode klasifikasi yang paling banyak digunakan dalam berbagai bidang penelitian karena akurasi yang cukup baik. Teknik Klasifikasi Naive Bayes untuk Prediksi Keterampilan Siswa menghasilkan akurasi 76,65%, yang mana lebih baik dibandingkan dengan metode Multilayer Perceptron yang hanya menghasilkan akurasi sebesar 71,2% serta Algoritma J48 yang menghasilkan akurasi sebesar 73,93% (Osmanbegović and Suljic, 2012). Algoritma Naive Bayes yang digunakan untuk Mencari Perkiraan Waktu Studi Mahasiswa menghasilkan tingkat kesalahan prediksi rata-rata sebesar 20% hingga 34% (Jananto, 2013). Penerapan Metode Naive Bayes pun dilakukan untuk memprediksi kelulusan mahasiswa dengan Data Mahasiswa Baru Perguruan Tinggi yang menghasilkan akurasi mencapai 94% (Syarli dan Muin, 2016).

Namun, meskipun memiliki akurasi yang cukup baik, sebenarnya akurasi dari Algoritma Naive Bayes masih dapat ditingkatkan dengan dilakukan seleksi fitur, seleksi fitur dilakukan untuk memilih atribut yang memiliki kontribusi paling banyak terhadap akurasi, karena Algoritma Naive Bayes menganggap semua atribut adalah independen/tidak saling terkait satu sama lain sehingga bobot kontribusi terhadap akurasi dianggap sama untuk setiap atribut (Wahyuni dan Luthfiarta, 2014). Independensi atribut pada metode Naive Bayes terjadi karena metode Naive Bayes menerapkan klasifikasi berdasarkan probabilitas, sehingga semua atribut perlu diperhitungkan secara individual untuk mendapatkan probabilitas data terhadap suatu kelas untuk menentukan hasil klasifikasi data tersebut (Kumar and Sahoo, 2015). Penelitian sebelumnya menunjukkan Algoritma Genetika merupakan salah satu metode yang dapat meningkatkan akurasi Algoritma Naive Bayes dengan

menggunakan seleksi fitur. Optimasi algoritma Naive Bayes dengan Algoritma Genetika yang melakukan seleksi fitur dapat meningkatkan akurasi sebesar 5,56% untuk Prediksi Hasil Pemilu Legislatif DKI Jakarta (Wahyuni dan Luthfiarta, 2014). Klasifikasi Penderita Penyakit Jantung dengan metode Naive Bayes yang dioptimasi dengan Algoritma Genetika berhasil meningkatkan 4% tingkat akurasi dari tingkat akurasi sebelum dilakukan optimasi (Kumar and Sahoo, 2015). Optimasi Algoritma Naive Bayes dengan Algoritma Genetika untuk prediksi kesuburan pun mengalami peningkatan akurasi dari 97,66% menjadi 99,33% (Buani, 2016).

Oleh karena itu, jika klasifikasi dilakukan dengan Algoritma Naive Bayes, masih ada kemungkinan selisih akurasi yang seharusnya bisa dicapai dengan optimasi, maka penelitian ini akan mengetahui bagaimana hasil optimasi Metode Naive Bayes dengan Algoritma Genetika untuk klasifikasi ketepatan waktu studi mahasiswa, khususnya mahasiswa Teknik Informatika Universitas Sriwijaya.

1.3 Rumusan Masalah

Berdasarkan latar belakang yang dikemukakan di atas, rumusan masalah pada penelitian ini adalah mengetahui bagaimana optimasi yang dapat dilakukan dengan algoritma genetika dalam memberikan pengaruh terhadap metode Naive Bayes untuk klasifikasi ketepatan waktu studi mahasiswa.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Mengetahui bagaimana optimasi yang dilakukan Algoritma Genetika dalam memberikan pengaruh terhadap metode Naive Bayes pada klasifikasi ketepatan waktu studi mahasiswa.
2. Mengetahui peningkatan akurasi yang terjadi setelah dilakukan optimasi dengan Algoritma Genetika terhadap klasifikasi ketepatan waktu studi mahasiswa dengan metode Naive Bayes.
3. Mengetahui jumlah atribut serta atribut apa saja yang dianggap penting sehingga berpengaruh terhadap optimasi yang dilakukan.

1.5 Manfaat Penelitian

Dengan melakukan penelitian ini, kita dapat mengetahui bagaimana optimasi yang dapat dilakukan oleh Algoritma Genetika terhadap klasifikasi ketepatan waktu studi mahasiswa menggunakan metode Naive Bayes.

1.6 Batasan Masalah

Adapun batasan dari penelitian ini adalah:

1. Data yang digunakan hanya berasal dari data mahasiswa Jurusan Teknik Informatika Kelas Reguler Universitas Sriwijaya dari angkatan 2011 hingga angkatan 2013.
2. Optimasi yang dimaksud pada penelitian ini adalah meningkatkan akurasi metode Naive Bayes dengan dilakukan seleksi fitur menggunakan Algoritma Genetika.

1.7 Sistematika Penulisan

Sistematika penulisan skripsi ini adalah sebagai berikut :

BAB I. PENDAHULUAN

Dalam bab ini akan diuraikan mengenai latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah atau ruang lingkup penelitian, serta sistematika penulisan.

BAB II. KAJIAN LITERATUR

Dalam bab ini akan diuraikan *literature review* yang dilakukan terhadap jurnal, buku, dan artikel, berkaitan dengan metode Naive Bayes, Algoritma Genetika, dan Optimasi Naive Bayes menggunakan Algoritma Genetika untuk meningkatkan akurasi.

BAB III. METODOLOGI PENELITIAN

Dalam bab ini akan diuraikan secara jelas tentang unit penelitian, metode pengumpulan data, tahapan penelitian, metode pengembangan perangkat lunak, serta manajemen proyek penelitian yang berupa penjadwalan penelitian.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Dalam bab ini akan diuraikan mengenai analisis dan perancangan perangkat lunak yang akan digunakan sebagai alat penelitian. Penjelasan dimulai dari pengumpulan dan analisa kebutuhan, rancangan, dan konstruksi perangkat lunak, serta pengujian yang bertujuan untuk memastikan semua kebutuhan

perangkat lunak sesuai dengan kebutuhan. Penyusunan pada bab ini memiliki kerangka penulisan dengan fase-fase dan elemen-elemen pengembangan perangkat lunak yang bersifat berorientasi objek.

BAB V. HASIL DAN ANALISA PENELITIAN

Dalam bab ini diuraikan hasil pengujian berdasarkan langkah-langkah yang telah direncanakan. Tabel hasil pengujian serta analisisnya disajikan sebagai dasar dari kesimpulan yang akan diambil dalam penelitian ini.

BAB VI. KESIMPULAN DAN SARAN

Dalam bab ini berisi kesimpulan dari semua uraian-uraian yang terdapat pada bab-bab sebelumnya dan juga saran-saran yang diharapkan berguna dalam pengaruh algoritma genetika untuk seleksi fitur terhadap klasifikasi Naive Bayes.

1.8 Kesimpulan

Dari pendahuluan ini, telah jelas diuraikan secara umum tentang penelitian yang dilakukan, meliputi latar belakang masalah, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah, serta sistematika penulisan. Adapun dijelaskan pula klasifikasi dalam bidang data mining dengan metode Naive Bayes memiliki akurasi cukup baik dibandingkan dengan metode lainnya, namun berdasarkan penelitian sebelumnya, akurasi Naive Bayes ini masih bisa ditingkatkan dengan dilakukannya optimasi menggunakan Algoritma Genetika.

Oleh karena itu, penelitian yang dilakukan akan mengetahui pengaruh dilakukannya optimasi metode Naive Bayes dengan Algoritma Genetika untuk klasifikasi ketepatan waktu studi mahasiswa, khususnya dalam penelitian ini, yang menjadi kasusnya adalah mahasiswa Jurusan Teknik Informatika Kelas Reguler Universitas Sriwijaya.

DAFTAR PUSTAKA

- Buani, D. C. P. (2016). Optimasi Algoritma Naive Bayes dengan Menggunakan Algoritma Genetika untuk Prediksi Kesuburan (Fertility). *Evolusi*, 4(1).
- Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier.
- Hastie, T. (2009). K-Fold Cross-Validation. *Sldm* 3, 3–11.
- Hermawanto, D. (2003). Algoritma Genetika dan contoh aplikasinya.
- Jananto, A. (2013). Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa. *Dinamik-Jurnal Teknologi Informasi*, 18(1).
- Kruchten, P. (2004). *The rational unified process: an introduction*. Addison-Wesley Professional.
- Kumar, S., & Sahoo, G. (2015). Classification of heart disease using Naive Bayes and genetic algorithm. In *Computational Intelligence in Data Mining-Volume 2* (pp. 269–282). Springer.
- Kusrini, E. T. L. (2009). Algoritma Data Mining. Yogyakarta: Andi Offset.
- Osmanbegović, E., & Suljic, M. (2012). Data mining approach for predicting student performance. *Economic Review*, 10(1), 3–12.
- Syarli, S., & Muin, A. A. (2016). Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi). *Jurnal Ilmu Komputer*, 2(1).
- Wahyuni, D. T., & Luthfiarta, A. (2014). PREDIKSI HASIL PEMILU LEGISLATIF DKI JAKARTA MENGGUNAKAN NAIVE BAYES DENGAN ALGORITMA GENETIKA SEBAGAI FITUR SELEKSI.