

**APLIKASI DATA MINING
MENGUNAKAN METODE NAIVE BAYES CLASSIFIER
STUDI KASUS KLASIFIKASI BEASISWA
DI UNIVERSITAS DIPONEGORO**



SKRIPSI

**Disusun Sebagai Salah Satu Syarat
untuk Memperoleh Gelar Sarjana Komputer
pada Jurusan Ilmu Komputer/ Informatika**

**Disusun oleh:
FAISAL RAHMAN
24010311130055**

**JURUSAN ILMU KOMPUTER/INFORMATIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO**

2015

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

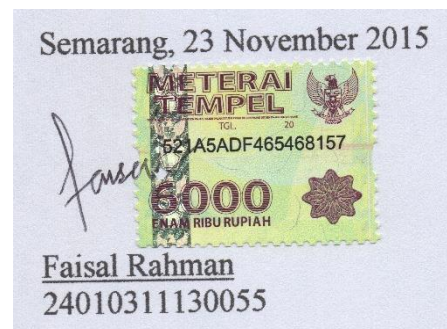
Saya yang bertanda tangan di bawah ini :

Nama : Faisal Rahman

NIM : 24010311130055

Judul : Aplikasi *Data Mining* Menggunakan Metode *Naive Bayes Classifier* Studi Kasus Klasifikasi Beasiswa di Universitas Diponegoro

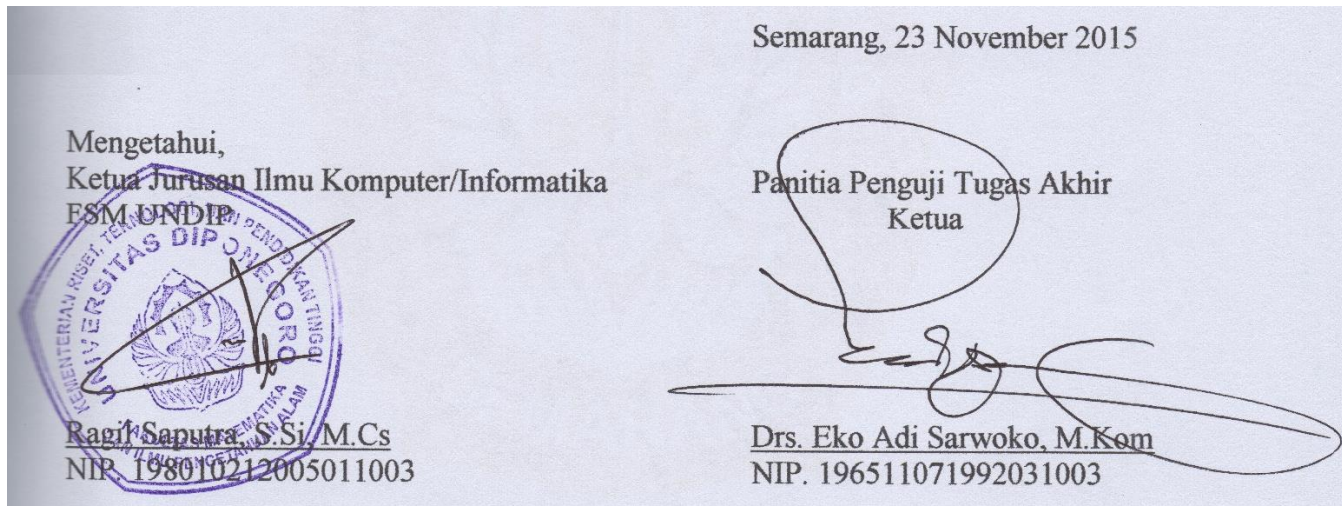
Dengan ini saya menyatakan bahwa dalam tugas akhir/ skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan di dalam daftar pustaka.



HALAMAN PENGESAHAN

Judul : Aplikasi *Data Mining* Menggunakan Metode *Naive Bayes Classifier* Studi Kasus Klasifikasi Beasiswa di Universitas Diponegoro
Nama : Faisal Rahman
NIM : 24010311130055

Telah diujikan pada sidang tugas akhir pada tanggal 2 November 2015 dan dinyatakan lulus pada tanggal 17 November 2015

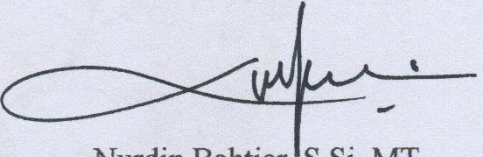


HALAMAN PENGESAHAN

Judul : Aplikasi *Data Mining* Menggunakan Metode *Naive Bayes Classifier* Studi Kasus Klasifikasi Beasiswa di Universitas Diponegoro
Nama : Faisal Rahman
NIM : 24010311130055

Telah diujikan pada sidang tugas akhir pada tanggal 2 November 2015.

Semarang, 23 November 2015
Dosen Pembimbing



Nurdin Bahtiar, S.Si, MT.
NIP. 197907202003121002

ABSTRAK

Universitas Diponegoro merupakan salah satu universitas di Indonesia yang menyediakan banyak beasiswa dari berbagai sumber. Beasiswa yang ditawarkan di Universitas Diponegoro bermacam-macam sehingga kriteria penerimaan dari setiap beasiswa berbeda-beda. Oleh karena itu, dibutuhkan aplikasi *data mining* klasifikasi beasiswa di Universitas Diponegoro yang dapat memberikan prediksi beasiswa yang cocok untuk mahasiswa Universitas Diponegoro. Aplikasi ini menggunakan metode naive bayes classifier dan dibangun dengan menggunakan bahasa pemrograman PHP dan basis data MySQL. Proses menghasilkan prediksi pada aplikasi ini yaitu dengan membandingkan atribut semester, IPK, pekerjaan ayah, pekerjaan ibu, penghasilan orangtua, dan jumlah tanggungan orangtua dari pengguna aplikasi dengan atribut yang sama dari pendaftar beasiswa yang ada pada basis data. Data yang digunakan pada aplikasi ini diambil dari Bagian Administrasi Kemahasiswaan Kesejahteraan Mahasiswa Universitas Diponegoro sebanyak 696 data yang dibagi ke dalam dua bagian yaitu 557 data latih dan 139 data uji. Aplikasi ini memiliki tingkat kesesuaian dengan data sebenarnya sebesar 89.93%.

Kata Kunci : Beasiswa, Universitas Diponegoro, *Data Mining*, *Naive Bayes Classifier*

ABSTRACT

Diponegoro University is one of the universities in Indonesia that provides many scholarships from various sources. Scholarships offered at University of Diponegoro manifold so that the acceptance criteria of each scholarship was different. Therefore, needed a University of Diponegoro scholarship classification data mining application that can a suitable prediction scholarships for students at the University of Diponegoro. This application uses naive bayes classifier method and built using the PHP programming language and MySQL database. The process of generating a prediction on this application is to compare the attributes of the semester, GPA, father's occupation, mother's occupation, parents' income and number of dependents of the parent application users with the same attributes of applicants scholarships that exist in the database. The data used in this application is taken from Bagian Administrasi Kemahasiswaan Kesejahteraan Mahasiswa Universitas Diponegoro as much as 696 data that divided into two sections, 557 training data and 139 test data. This application has a degree of conformity with the actual data of 89.93%.

Keywords : Scholarship, University of Diponegoro, Data Mining, Naive Bayes classifier

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan laporan tugas akhir yang berjudul “Aplikasi *Data Mining* Menggunakan Metode Naive Bayes Classifier Studi Kasus Klasifikasi Beasiswa Di Universitas Diponegoro” dengan baik dan lancar. Laporan tugas akhir ini disusun sebagai salah satu syarat untuk memperoleh gelar sarjana strata satu pada Jurusan Ilmu Komputer/ Informatika Fakultas Sains dan Matematika Universitas Diponegoro Semarang.

Dalam penyusunan tugas akhir ini penulis banyak mendapat bimbingan, bantuan, dan dukungan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis menyampaikan terimakasih kepada:

1. Prof. Dr. Widowati, M.Si, selaku Dekan FSM UNDIP
2. Ragil Saputra, S.Si, M.Cs, selaku Ketua Jurusan Ilmu Komputer/ Informatika
3. Helmie Arif Wibawa, S.Si, M.Cs, selaku Koordinator Tugas Akhir
4. Nurdin Bahtiar, S.Si, M.T, selaku dosen pembimbing

Penulis menyadari bahwa dalam laporan ini masih banyak kekurangan baik dari penyampaian materi maupun isi dari materi itu sendiri. Hal ini dikarenakan keterbatasan kemampuan dan pengetahuan dari penulis. Oleh karena itu, kritik dan saran yang bersifat membangun sangat penulis harapkan. Semoga laporan tugas akhir ini dapat bermanfaat bagi penulis dan juga pembaca pada umumnya.

Semarang, 23 November 2015

Faisal Rahman

DAFTAR ISI

| | |
|---|------|
| HALAMAN PERNYATAAN KEASLIAN SKRIPSI..... | ii |
| HALAMAN PENGESAHAN | iii |
| HALAMAN PENGESAHAN | iv |
| ABSTRAK | v |
| <i>ABSTRACT</i> | vi |
| KATA PENGANTAR..... | vii |
| DAFTAR ISI | viii |
| DAFTAR GAMBAR..... | x |
| DAFTAR TABEL | xii |
| LAMPIRAN | xiii |
| BAB I PENDAHULUAN | 1 |
| 1.1. Latar Belakang | 1 |
| 1.2. Rumusan Masalah | 2 |
| 1.3. Tujuan dan Manfaat..... | 2 |
| 1.4. Ruang Lingkup | 3 |
| 1.5. Sistematika Penulisan..... | 3 |
| BAB II TINJAUAN PUSTAKA | 5 |
| 2.1. <i>Data Mining</i> | 5 |
| 2.2. <i>Arsitektur Data Mining</i> | 6 |
| 2.3. Tahap-tahap <i>Data Mining</i> | 7 |
| 2.4. Teknik-teknik <i>Data Mining</i> | 9 |
| 2.5. <i>Naive Bayes Classifier</i> | 10 |
| 2.6. <i>Laplacian Correction</i> | 11 |
| 2.7. <i>Nilai Support</i> | 11 |
| 2.8. Bahasa Pemrograman PHP..... | 11 |
| 2.9. <i>Database Management System</i> | 12 |
| 2.10. <i>Database Management System MySQL</i> | 14 |
| 2.11. Pemodelan Data..... | 15 |
| 2.12. Pemodelan Fungsional | 15 |
| 2.13. Perancangan Perangkat Lunak | 16 |
| 2.14. Integrasi dan Pengujian Sistem | 17 |

| | |
|--|----|
| BAB III ANALISIS DAN PERANCANGAN | 18 |
| 3.1. Analisis Kebutuhan | 18 |
| 3.1.1. Perspektif Aplikasi yang Akan Dikembangkan | 18 |
| 3.1.2. Tahap-tahap Pembuatan Aplikasi <i>Data Mining</i> | 18 |
| 3.2. Analisis Aplikasi yang Dikembangkan | 27 |
| 3.2.1. Spesifikasi Kebutuhan Fungsional | 27 |
| 3.2.2. Pemodelan Data | 27 |
| 3.2.2.1. Kamus Data (Data Dictionary) | 28 |
| 3.2.3. Pemodelan Fungsional | 29 |
| 3.2.3.1. DFD Level 0 / DCD (Data Context Diagram) | 30 |
| 3.2.3.2. DFD Level 1 | 30 |
| 3.2.3.3. DFD Level 2 Proses 2 | 31 |
| 3.2.3.4. DFD Level 2 Proses 5 | 32 |
| 3.3. Kebutuhan Antarmuka | 33 |
| 3.3.1. Kebutuhan Antarmuka Pengguna | 33 |
| 3.3.2. Kebutuhan Antarmuka Perangkat Lunak | 33 |
| 3.3.3. Kebutuhan Antarmuka Perangkat Keras | 33 |
| 3.4. Perancangan | 34 |
| 3.4.1. Perancangan Fungsi | 34 |
| 3.4.2. Perancangan Antarmuka | 42 |
| BAB IV IMPLEMENTASI DAN PENGUJIAN | 52 |
| 4.1. Implementasi | 52 |
| 4.1.1. Implementasi Perangkat Keras dan Perangkat Lunak | 52 |
| 4.1.2. Implementasi Perancangan Data | 52 |
| 4.1.3. Implementasi Perancangan Fungsi | 53 |
| 4.1.4. Implementasi Perancangan Antarmuka | 60 |
| 4.2. Pengujian | 68 |
| 4.2.1. Pengujian Fungsi-Fungsi Aplikasi | 68 |
| 4.2.1.1. Lingkungan Pengujian | 69 |
| 4.2.1.2. Identifikasi dan Rencana Pengujian | 69 |
| 4.2.1.3. Hasil Uji | 70 |
| 4.2.2. Pengujian Hasil Aplikasi | 70 |
| BAB V PENUTUP | 72 |
| 5.1. Kesimpulan | 72 |
| 5.2. Saran | 72 |
| DAFTAR PUSTAKA | 73 |
| LAMPIRAN-LAMPIRAN | 75 |

DAFTAR GAMBAR

| | |
|--|----|
| Gambar 2.1. Arsitektur <i>Data Mining</i> (Han & Kamber, 2006) | 6 |
| Gambar 2.2. Proses <i>Data Mining</i> (Han & Kamber, 2006)..... | 8 |
| Gambar 3.1 DFD Level 0 | 30 |
| Gambar 3.2 DFD Level 1 | 30 |
| Gambar 3.3 DFD Level 2 Proses 2..... | 31 |
| Gambar 3.4 DFD Level 2 Proses 5..... | 32 |
| Gambar 3.5 Antarmuka Awal..... | 43 |
| Gambar 3.6 Antarmuka Jenis Beasiswa yang Ada di Universitas Diponegoro | 43 |
| Gambar 3.7 Antarmuka Tentang Aplikasi..... | 44 |
| Gambar 3.8 Antarmuka Form <i>Input Aplikasi Data Mining</i> | 45 |
| Gambar 3.9 Antarmuka Hasil Proses <i>Data Mining</i> | 45 |
| Gambar 3.10 Antarmuka <i>Login Admin</i> | 46 |
| Gambar 3.11 Antarmuka Data Mahasiswa Pendaftar Beasiswa | 46 |
| Gambar 3.12 Antarmuka Ganti <i>Password</i> | 48 |
| Gambar 3.13 Antarmuka Tambah Fakultas..... | 48 |
| Gambar 3.14 Antarmuka Tambah Jurusan | 49 |
| Gambar 3.15 Antarmuka Tambah Data..... | 49 |
| Gambar 3.16 Antarmuka Tambah Data Berhasil | 49 |
| Gambar 3.17 Antarmuka <i>Import Data</i> | 49 |
| Gambar 3.18 Antarmuka <i>Import Data</i> Berhasil..... | 50 |
| Gambar 3.19 Antarmuka <i>Edit Data</i> | 50 |
| Gambar 3.20 Antarmuka <i>Edit Data</i> Berhasil..... | 51 |
| Gambar 4.1 Antarmuka Awal..... | 61 |
| Gambar 4.2 Antarmuka Jenis Beasiswa | 61 |
| Gambar 4.3 Antarmuka Tentang Aplikasi..... | 62 |
| Gambar 4.4 Antarmuka <i>Form Input Aplikasi Data Mining</i> | 62 |
| Gambar 4.5 Antarmuka Hasil Proses <i>Data Mining</i> | 63 |
| Gambar 4.6 Antarmuka <i>Login Admin</i> | 63 |
| Gambar 4.7 Antarmuka Data Mahasiswa Pendaftar Beasiswa | 64 |
| Gambar 4.8 Antarmuka Ganti <i>Password</i> | 65 |
| Gambar 4.9 Antarmuka Tambah Fakultas..... | 66 |

| | |
|--|----|
| Gambar 4.10 Antarmuka Tambah Jurusan | 66 |
| Gambar 4.11 Antarmuka Tambah Data..... | 66 |
| Gambar 4.12 Antarmuka Tambah Data Berhasil | 66 |
| Gambar 4.13 Antarmuka <i>Import</i> Data..... | 67 |
| Gambar 4.14 Antarmuka <i>Import</i> Data Berhasil..... | 67 |
| Gambar 4.15 Antarmuka <i>Edit</i> Data | 68 |
| Gambar 4.16 Antarmuka <i>Edit</i> Data Berhasil..... | 68 |

DAFTAR TABEL

| | |
|---|----|
| Tabel 2.1 Notasi <i>Data Dictionary</i> (Pressman, 2002)..... | 15 |
| Tabel 2.2 Notasi DFD (Pressman, 2002)..... | 16 |
| Tabel 3.1 Tabel <i>System Requirement Specification</i> | 28 |
| Tabel 4.1 Struktur Tabel Beasiswa..... | 54 |
| Tabel 4.2 Struktur Tabel <i>Admin</i> | 54 |
| Tabel 4.3 Identifikasi dan Rencana Pengujian..... | 70 |

LAMPIRAN

| | |
|---|-----|
| Lampiran 1. Hasil Uji..... | 76 |
| Lampiran 2. Data Latih Aplikasi..... | 83 |
| Lampiran 3. Pengujian Hasil Aplikasi..... | 108 |
| Lampiran 4. Data Contoh Perhitungan Algoritma Naive Bayes Classifier..... | 115 |

BAB I

PENDAHULUAN

Bab pendahuluan menguraikan tentang latar belakang, rumusan masalah, tujuan dan manfaat, ruang lingkup, dan sistematika penulisan dari aplikasi *data mining* menggunakan metode *naive bayes classifier* studi kasus klasifikasi beasiswa di Universitas Diponegoro.

1.1. Latar Belakang

Dewasa ini kemajuan teknologi informasi dan komputer telah menyediakan fasilitas penyimpanan data yang besar dalam *format* elektronik sehingga penyimpanan data bukan lagi menjadi satu hal yang sulit. Sebagai konsekuensinya jumlah data yang disimpan mengalami peningkatan yang sangat cepat dari segi kuantitas dan kualitas.

Pada beberapa perguruan tinggi besar, data dapat diperoleh dari data historis dan data kegiatan operasional sebuah perguruan tinggi, dimana data ini akan bertambah secara terus menerus, sehingga proses eksplorasi data dalam menentukan hubungan antar variabel di dalam data menjadi sangat lambat dan memiliki proses yang subjektif. Salah satu solusi yang mungkin digunakan untuk menangani masalah ini adalah *data mining*.

Alasan utama mengapa *data mining* sangat menarik perhatian industri informasi dalam beberapa tahun belakangan ini adalah karena tersedianya data dalam jumlah yang besar dan semakin besarnya kebutuhan untuk mengubah data tersebut menjadi informasi dan pengetahuan yang berguna (Fadli, 2011).

Universitas Diponegoro yang merupakan salah satu perguruan tinggi besar di Indonesia sangat membutuhkan *data mining* untuk mengolah data yang dimiliki menjadi informasi yang berguna. Sebagai salah satu universitas terbesar di Indonesia banyak beasiswa yang ditawarkan berbagai instansi untuk mahasiswa-mahasiswa Universitas Diponegoro. Banyak beasiswa yang ada di Universitas Diponegoro membuat mahasiswa dapat memilih beasiswa sesuai dengan kapasitas mereka. Tetapi mahasiswa mengalami kesulitan untuk menentukan beasiswa yang sesuai.

Kriteria penerimaan beasiswa dari setiap instansi berbeda-beda, baik itu dari atribut-atribut yang dipilih sebagai penilaian maupun standarisasi penilaian

tersebut. Untuk itu diperlukan *data mining* untuk melakukan analisis kecocokan data mahasiswa yang ingin mengajukan beasiswa dengan data mahasiswa yang telah menerima beasiswa dari berbagai instansi sehingga didapatkan beasiswa yang cocok untuk mahasiswa yang ingin mengajukan beasiswa.

Teknik *data mining* klasifikasi sesuai untuk diterapkan dalam kasus ini. Klasifikasi adalah teknik *data mining* yang menggunakan satu set data yang telah terklasifikasi sebagai perbandingan untuk mengembangkan model yang dapat mengklasifikasikan seluruh data yang ada (Ramageri, 2010).

Klasifikasi mengacu pada tugas memprediksi kelas suatu data untuk data yang belum mempunyai kelas (Zaki & Wagner JR, 2014). Ada beberapa algoritma *data mining* untuk klasifikasi dan salah satunya adalah *naive bayes classifier*. Algoritma *naive bayes classifier* digunakan pada pembuatan aplikasi ini karena *naive bayes classifier* hanya membutuhkan sejumlah kecil data pelatihan untuk memperkirakan parameter yang diperlukan. Pelatihan dan pengklasifikasian *naive bayes classifier* dapat sangat cepat dibandingkan dengan metode lain (Zhang, 2004). *Naive bayes classifier* mengasumsikan bahwa semua atribut independen. Hal ini menjadikannya sederhana tetapi efektif dalam melakukan klasifikasi (Zaki & Wagner JR, 2014).

Berdasarkan uraian diatas maka akan dibangun sebuah aplikasi *data mining* menggunakan metode *naive bayes classifier* studi kasus klasifikasi beasiswa di Universitas Diponegoro.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka dapat dirumuskan masalah yaitu bagaimana membuat aplikasi *data mining* menggunakan metode *naive bayes classifier* studi kasus klasifikasi beasiswa di Universitas Diponegoro.

1.3. Tujuan dan Manfaat

Hal-hal yang menjadi tujuan dari tugas akhir ini adalah:

1. Menghasilkan aplikasi yang dapat melakukan klasifikasi terhadap jenis beasiswa yang terdapat di Universitas Diponegoro.

Manfaat dari tugas akhir ini adalah:

1. Bagi penulis

Mengimplementasikan secara langsung materi yang diperoleh selama berada di perkuliahan, terutama dalam hal bidang *data mining*.

2. Bagi mahasiswa Universitas Diponegoro

Menjadi bahan pertimbangan dalam memilih jenis beasiswa yang sesuai.

3. Bagi Universitas Diponegoro

Memudahkan dalam proses seleksi beasiswa karena mahasiswa dapat memilih pilihan beasiswa yang sesuai saja tidak mendaftar pada setiap jenis beasiswa.

1.4. Ruang Lingkup

Ruang lingkup dari pengembangan aplikasi *data mining* ini dapat bermanfaat dalam menentukan pilihan beasiswa untuk mahasiswa Universitas Diponegoro adalah sebagai berikut:

1. Data diambil dari Bagian Administrasi Kemahasiswaan Kesejahteraan Mahasiswa Universitas Diponegoro.
2. Data terdiri dari 18 atribut yang terdiri dari 696 data mahasiswa Universitas Diponegoro.
3. Klasifikasi dilakukan pada lima jenis beasiswa diantaranya beasiswa Sobat Bumi Pertamina, beasiswa PT. Angkasa Pura (BUMN), beasiswa Marga Jaya Sejahtera, beasiswa PPA dan beasiswa Bank Indonesia.
4. Pembangunan aplikasi menggunakan Bahasa Pemrograman PHP dan *Database Management System MySQL*.

1.5. Sistematika Penulisan

Sistematika penulisan yang digunakan dalam tugas akhir ini terbagi dalam beberapa pokok bahasan, yaitu :

BAB I PENDAHULUAN

Berisi tentang latar belakang masalah, rumusan masalah, tujuan dan manfaat penulisan tugas akhir, ruang lingkup masalah, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Berisi kumpulan studi pustaka yang berhubungan dengan topik tugas akhir.

BAB III ANALISIS DAN PERANCANGAN

Bab ini membahas analisis dan perancangan aplikasi *data mining*.

BAB IV IMPLEMENTASI DAN PENGUJIAN

Bab ini berisi penerapan teknik *data mining*, pembuatan *prototype* aplikasi *data mining* dan pengujian.

BAB V PENUTUP

Bab ini berisi kesimpulan yang diambil berkaitan dengan aplikasi *data mining* dan saran-saran kepada semua pihak yang bersangkutan.