

## PERNYATAAN

Saya sebagai penulis menyatakan dengan sesungguhnya bahwa penelitian ini adalah hasil penelitian asli dari diri saya sendiri. Jika terdapat karya orang lain atau pun referensi dari karya orang lain saya akan mencantumkan sumber dengan jelas

Demikian pernyataan ini saya buat dengan keadaan sadar tanpa paksaan dari pihak manapun.

Yogyakarta, 06 Februari 2017

Penulis,



Arya Mardhani

20120140053

## PRAKATA

Puji syukur dengan kehadiran Allah SWT yang telah melimpahkan rahmat dan hidayahnya, sehingga penulis dapat menyelesaikan tugas akhir yang berjudul “PENERAPAN *DATA MINING* DALAM PREDIKSI KELULUSAN TEPAT WAKTU MENGGUNAKAN ALGORITMA *NAIVE BAYES*”. Laporan skripsi ini disusun untuk memenuhi salah satu syarat dalam memperoleh gelar Sarjana Teknik (ST) pada Program Studi S1 Teknik Informatika Universitas Muhammadiyah Yogyakarta.

Penulis menyadari bahwasanya laporan skripsi tidak dapat terselesaikan tanpa bantuan dan dukungan dari berbagai pihak. Penulis mengucapkan terima kasih yang tak terhingga kepada:

1. Bapak Ir. Eko Prasetyo, M.Eng., selaku pembimbing utama yang dengan penuh kesabaran memberikan ilmu, masukan, bimbingan, bantuan dan pengetahuan kepada penulis baik dalam hal teknis atau non teknis, selama penulisan skripsi maupun selama masa perkuliahan penulis dan atas setiap kepercayaan yang diberikan.
2. Bapak Slamet Riyadi, S.T., M.Sc., Ph.D., selaku dosen pembimbing pendamping yang telah meluangkan waktunya dan sangat sabar membimbing penulis dalam pembuatan skripsi.
3. Ibu Aprilia Kurnianti, S.T., M.Eng., selaku dosen penguji yang mau meluangkan waktu untuk memberikan kritik dan saran kepada penulis.

4. Bapak Muhammad Helmi Zain, S.T., M.T., selaku Ketua Program Studi S1 Teknik Informatika Universitas Muhammadiyah Yogyakarta, yang telah dengan tulus berbagi ilmu dan pengalaman selama perkuliahan penulis.
5. Para Dosen dan Staff Teknik Informatika yang senantiasa dengan ikhlas memberikan ilmu dan berbagi cerita atau pengalaman yang bermanfaat bagi penulis.
6. Dede, Yudi, Agung yang telah membantu memberikan masukan dan saran selama pembuatan skripsi ini sehingga penulis dapat menyelesaikannya.
7. Teman teman Teknik Informatika khususnya teman seperjuangan angkatan 2012 B sungguh mengenal kalian sebagai sahabat yang bersama-sama meraih impian dan asa adalah suatu hal terindah yang dikirimkan oleh Allah SWT dalam hidup penulis.
8. Terimakasih Elvan diano, maga ringga, Suryatman, Dwiki, Hidayatul, Aditya Herwanto, yang telah memberi semangat penulis dalam mengerjakan skripsi.
9. Kakak tingkat 2010, 2011, adik tingkat 2013, 2014, 2015 dan anak kos ijo yang tidak bisa penulis cantumkan satu persatu, terimakasih banyak kepada kalian yang selalu menemani penulis.

## DAFTAR ISI

HALAMAN PENGESAHAN I .....	II
PENERAPAN <i>DATA MINING</i> DALAM PREDIKSI KELULUSAN TEPAT WAKTU MENGGUNAKAN ALGORITMA <i>NAIVE BAYES</i> .....	II
HALAMAN PENGESAHAN II .....	III
PERNYATAAN.....	IV
PRAKATA.....	V
DAFTAR ISI.....	VII
DAFTAR GAMBAR .....	X
DAFTAR TABEL.....	XI
ABSTRACT .....	XII
INTISARI.....	XIII
BAB I .....	1
PENDAHULUAN .....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	1
1.3. Tujuan Penelitian.....	2
1.4. Batasan Masalah .....	2
1.5. Manfaat Penelitian .....	2
1.6. Sistematika Penulisan.....	2
BAB II.....	5
TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	5
2.1. Tinjauan Pustaka .....	5
2.2. Landasan Teori.....	7
2.2.1. <i>Data Mining</i> .....	7
2.2.2. <i>Pengelompokan Data Mining</i> .....	7
2.2.3 <i>Tahapan-tahapan Data Mining</i> .....	9
2.2.4. <i>Klasifikasi</i> .....	10
2.2.5. <i>Indeks Prestasi Kumulatif (IPK)</i> .....	11

2.2.6. <i>Naive Bayes</i> .....	11
2.2.7. <i>Rapidminer</i> .....	14
2.2.8. <i>Microsoft SQL</i> .....	16
2.2.9. <i>Microsoft Excel</i> .....	16
2.2.10. <i>Cross Validation</i> .....	17
BAB III .....	18
METODOLOGI PENELITIAN.....	18
3.1. Tempat dan waktu Penelitian.....	18
3.2. Alat dan Bahan .....	18
3.2.1 Alat Penelitian.....	18
3.3. Alur Penelitian.....	19
3.5 Pengumpulan Data .....	20
3.6 Seleksi Data.....	20
3.7 Pembersihan Data.....	20
3.8 Transformasi Data.....	21
3.9 Implementasi .....	21
4.0. Analisis Data .....	21
BAB IV .....	23
HASIL DAN PEMBAHASAN.....	23
4.1. Pengumpulan Data .....	23
4.2. Seleksi Data.....	25
4.3. Pembersihan Data.....	25
4.4. Tranfosrmasi Data.....	25
4.5. Implementasi .....	28
4.6. Pengujian <i>software Rapidminer</i> .....	30
4.7 Algoritma <i>Naïve Bayes</i> .....	46
BAB V.....	52
KESIMPULAN DAN SARAN.....	52
5.1 Kesimpulan .....	52
5.2 Saran.....	52
DAFTAR PUSTAKA .....	54

LAMPIRAN .....	55
<i>Perhitungan Manual</i> .....	81

## DAFTAR GAMBAR

Gambar 2. 1 Proses klasifikasi .....	10
Gambar 2. 2 Alur Metode <i>Naive Bayes</i> .....	13
Gambar 2. 3 <i>K-fold Validation</i> .....	17
Gambar 3. 1 <i>Flowchart</i> alur penelitian .....	19
Gambar 4. 2 <i>View Data</i> .....	24
Gambar 4. 1 <i>Connect Server</i> .....	23
Gambar 4. 2 <i>View Data</i> .....	24
Gambar 4. 3 Transformasi Data .....	26
Gambar 4. 4 Data yang telah diinisialisasi.....	28
Gambar 4. 5 <i>Data Training</i> format <i>csv</i> .....	29
Gambar 4. 6 <i>Data Training</i> format <i>csv</i> .....	29
Gambar 4. 7 <i>Drag and drop csv</i> .....	30
Gambar 4. 8 <i>Importing file training</i> .....	31
Gambar 4. 9 <i>Form data import data training</i> .....	31
Gambar 4. 10 <i>Data importing wizard step 2 data training</i> .....	32
Gambar 4. 11 <i>Data importing wizard step 3 data training</i> .....	33
Gambar 4. 12 <i>Data importing wizard step 4 data training</i> .....	33
Gambar 4. 13 <i>Importing file csv testing</i> .....	34
Gambar 4. 14 <i>Form data import data testing</i> .....	35
Gambar 4. 15 <i>Data importing wizard step 2 data training</i> .....	35
Gambar 4. 16 <i>Data importing wizard step 3 data testing</i> .....	36
Gambar 4. 17 <i>Data importing wizard step 4 data testing</i> .....	37
Gambar 4. 18 <i>File read csv training dengan operator naïve bayes</i> .....	38
Gambar 4. 19 <i>Apply model</i> .....	38
Gambar 4. 20 <i>Operator performance</i> .....	39
Gambar 4. 21 Menghubungkan <i>operator performance</i> .....	40
Gambar 4. 22 Menjalankan <i>Rapidminer</i> .....	40
Gambar 4. 23 Hasil perhitungan .....	41
Gambar 4. 24 <i>Accuracy</i> .....	42
Gambar 4. 25 <i>Grafik Scatter</i> .....	42
Gambar 4. 26 <i>Cross validation</i> .....	43
Gambar 4. 27 <i>Cross validation view</i> .....	44
Gambar 4. 28 <i>Parameters cross validation</i> .....	44
Gambar 4. 29 Menjalankan <i>rapidminer</i> .....	45
Gambar 4. 30 Hasil <i>shuffled sampling</i> .....	45

## DAFTAR TABEL

Tabel 1 Inialisasi nilai IPK .....	27
Tabel 2 Inialisasi asal pulau .....	27
Tabel 3 Perbandingan akurasi .....	46
Tabel 4 <i>Data training</i> mahasiswa angkatan jurusan HI 2010 – 2011 .....	46
Tabel 5 <i>Data testing</i> mahasiswa angkatan jurusan HI 2011 – 2014 .....	47
Tabel 6 <i>Data Training</i> lengkap .....	55
Tabel 7 <i>Data Training</i> lengkap .....	56