

## PERNYATAAN

Saya menyatakan bahwa skripsi ini adalah benar karya saya sendiri dan bukan plagiat dari orang lain. Apabila di kemudian hari terbukti melakukan plagiat dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan peraturan perundang-undangan.

Bojonegoro, 31 Juli 2023



Mariyatul Qibtiyah

NIM : 2520190054

## HALAMAN PERSETUJUAN

Nama : Mariyatul Qibtiyah

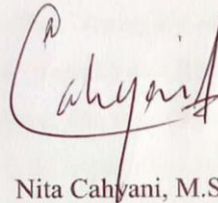
NIM : 2520190054

Judul : Prediksi Tingkat Kelahiran Bayi di Kabupaten Bojonegoro dengan  
Menggunakan Algoritma Naive Bayes

Telah disetujui dan dinyatakan memenuhi syarat untuk diajukan dalam ujian  
Skripsi.

Bojonegoro, 09 Agustus 2023.

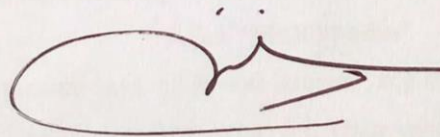
Pembimbing I



Nita Cahyani, M.Stat.

NIDN.0704038906

Pembimbing II



Fetrika Anggraini, M.Pd.

NIDN.0718038803

## HALAMAN PENGESAHAN

Nama : Mariyatul Qibtiyah

NIM : 2520190054

Judul : Prediksi Tingkat Kelahiran Bayi di Kabupaten Bojonegoro dengan  
Menggunakan Algoritma Naive Bayes

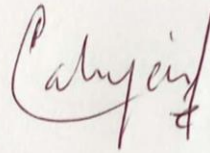
Telah dipertahankan di hadapan penguji pada tanggal 21 Agustus 2023.

Dewan Penguji

Tim Pembimbing

Penguji I

Pembimbing I



Alif Yuanita Kartini, M.Si

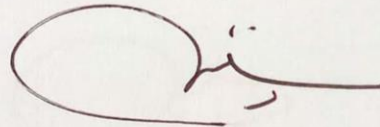
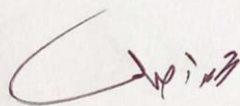
Nita Cahyani, M.Stat

NIDN: 0721048606

NIDN:0704038906

Penguji II

Pembimbing II



Dr.H.M. Ridlwan Hambali, Lc., M.A.

Fetrika Anggraini, M.Pd

NIDN:2117056803

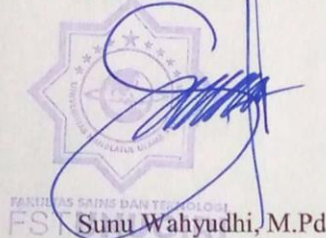
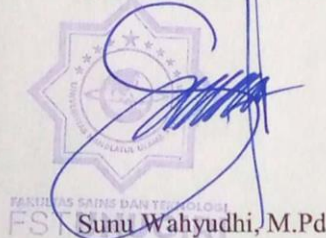
NIDN:0718038803

Mengetahui,

Mengetahui,

Dekan Fakultas Sains dan Teknologi

Ketua Program Studi

Sunu Wahyudhi, M.Pd

NIDN:0709058902


Nita Cahyani, M.Stat

NIDN:0704038906

## MOTTO

“Allah tidak berjanji bahwa langit akan selalu biru, tetapi Allah berjanji bersama kesulitan pasti ada kemudahan”

Tidak ada sesuatu yang mustahil untuk dicapai. Tidak ada sesuatu yang mustahil untuk diselesaikan. Karena “*Sesungguhnya Allah bebas melaksanakan kehendak-Nya, Dia telah menjadikan untuk setiap sesuatu menurut takarannya.*”

(QS. At-Thalaq:3)

“Berusahalah untuk tidak menjadi manusia yang berhasil, tetapi berusahalah menjadi manusia yang berguna”

(Albert Einstein)

## PERSEMBAHAN

**Kedua Orang tua tercinta**

**Nur Qo'im (Bapak) dan Munawaroh (Ibu)**

Sebagai tanda bakti, hormat, dan rasa terima kasih yang tiada terhingga kupersembahkan karya kecil ini kepada Bapak dan Ibu yang telah memberikan kasih sayang, support system, dukungan moril maupun materi serta doa yang tiada hentinya, dan cinta kasih yang tiada terhingga yang tiada mungkin dapat kubalas hanya dengan selembar kertas yang bertuliskan cinta dalam kata persembahan. Tiada kata seindah lantunan doa dan tiada doa yang paling khusuk selain doa yang terucap dari orang tua. Semoga engkau diberikan kesehatan dan panjang umur agar dapat menemani di setiap langkah kecilku bersama adek menuju kesuksesan.

### **Adek “Muhibbudin”**

Satu-satunya saudara saya, adek saya tercinta yang selalu memberikan doa dan semangat untuk menyelesaikan skripsi ini. Tiada yang paling mengharukan saat kumpul bersamamu, walaupun sering bertengkar tetapi hal itu selalu menjadi warna yang tak akan bisa tergantikan, hanya karya kecil ini yang dapat aku persembahkan. Tuntutlah ilmu setinggi mungkin, dan harumkan nama baik keluarga, maaf jika belum bisa menjadi panutan seutuhnya, tetapi aku akan selalu menjadi yang terbaik untuk kamu dek.

## KATA PENGANTAR

Segala puji dan syukur atas kehadiran Allah SWT, karena atas segala kenikmatan dan kekuatannya sehingga penulis dapat menyelesaikan penyusunan proposal skripsi yang berjudul **“Prediksi Tingkat Kelahiran Bayi di Kabupaten Bojonegoro dengan Menggunakan Algoritma Naive Bayes”**. Penyusunan proposal skripsi ini sebagai salah satu syarat mencapai gelar Strata-1 pada program studi Statistika Fakultas Sains dan Teknologi, Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.

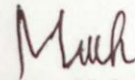
Dalam proses pelaksanaan dan pembuatan laporan ini, penulis mendapatkan dukungan dan bantuan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada :

1. Bapak KH M. Jauharul Ma'arif, M. Pd. I selaku Rektor Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
2. Bapak Sunu Wahyudhi, M.Pd. selaku Dekan Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
3. Ibu Nita Cahyani, M.Stat. selaku Ketua Program Studi Statistika Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
4. Ibu Nita Cahyani, M.Stat. selaku dosen pembimbing 1 yang ditengah segala kesibukannya dapat meluangkan waktu untuk memberikan bimbingan, saran, masukan, serta motivasi selama penyusunan laporan ini.
5. Ibu Fetrika Anggraini, M.Pd. selaku dosen pembimbing 2 yang telah membantu dalam penulisan laporan skripsi ini.
6. Seluruh dosen program studi Statistika yang telah memberikan ilmu dan pengalaman yang bermanfaat kepada penulis.
7. Kedua orang tua dan keluarga yang telah memberikan semangat dan senantiasa mendukung penulis baik secara moril maupun materil.
8. Teman-teman seperjuangan untuk segala dukungan, bantuan, keceriaan dan kebersamaannya selama ini.
9. Teman-teman prodi Statistika 2019 yang tidak bisa disebutkan satu-persatu.
10. Seluruh pihak yang secara langsung maupun tidak langsung yang telah membantu penulis dalam menyelesaikan proposal skripsi ini.

Penulis menyadari akan keterbatasan dan kekurangan dalam penyusunan skripsi. Oleh karena itu, kritik dan saran sangat penulis harapkan agar kedepannya dapat menjadi lebih baik. Semoga laporan ini dapat membawa manfaat bagi semua pihak, umumnya bagi pembaca dan khususnya bagi penulis.

Bojonegoro, 20 Februari 2023

Penulis



Mariyatul Qibtiyah



# UNUGIRI

## ABSTRACT

*Qibtiyah, Mariyatul. 2023. Prediction of Birth Rate in Bojonegoro District Using The Naive Bayes Algorithm. Thesis, Department of Statistics, Faculty of Science and Technology, Nahdlatul Ulama Sunan Giri University. Main Advisor Nita Cahyani, M.Stat. and Advisor Fetrika Anggraini, M.Pd. The development of information technology in the times this is very fast starting from the discovery of new information on big data with looking for a certain pattern or often referred to as Data Mining. Data Mining is a process which uses statistical techniques, mathematics, artificial intelligence, as well as machine learning used to extract (separation process) and identify related useful information from large databases. With the current development of information technology, especially in the field of data mining which has been widely used in the system information. There are several data mining classification methods that are often used to predict or forecast, such as the Naive Bayes Classifier Algorithm, Decision Tree, Neural Network, K-Nearest Neighbor, Artificial Neural Network, and other classification methods and so on. The method used in the research it uses the Naive Bayes Algorithm method that utilizes the method probability and statistics put forward by the British scientist Thomas Bayes, namely scientists who predict future probabilities based on previous or past experiences. Influencing factors The birth rate itself is a factor of male sex, sex women, and couples of childbearing age. Baby birth rate as an important indicator to reflect the state of health degree in a society. The data used from 1993 to 2022. The results of the Naive Bayes Algorithm analysis in this study obtained that 27 data with low predictions and 3 data with high predictions in 2017, 2020, 2021 and an accuracy value (accuracy) Recall and Precision of 100%.*

**Keywords :** *Data Mining, Prediction, Naive Bayes Algorithm*

UNUGIRI

## ABSTRAK

Qibtiyah, Mariyatul. 2023. *Prediksi Tingkat Kelahiran Bayi di Kabupaten Bojonegoro dengan Menggunakan Algoritma Naive Bayes*. Skripsi. Program Studi Statistika Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri. Pembimbing Utama Nita Cahyani, M.Stat. dan Pembimbing Pendamping Fetrika Anggraini, M.Pd. Perkembangan teknologi informasi pada zaman ini yang sangat pesat dimulai dari penemuan informasi baru pada *big data* dengan mencari suatu pola tertentu atau sering disebut dengan istilah *Data Mining*. *Data mining* merupakan proses dimana menggunakan teknik statistik, matematika, kecerdasan buatan, serta *machine learning* yang digunakan untuk mengekstraksi (proses pemisahan) dan mengidentifikasi informasi yang bermanfaat terkait dari berbagai database besar. Dengan berkembangnya teknologi informasi saat ini, terutama pada bidang *data mining* yang telah banyak digunakan dalam sistem informasi. Ada beberapa metode klasifikasi *data mining* yang sering digunakan untuk memprediksi atau meramalkan, seperti *Algoritma Naive Bayes Classifier*, *Decision Tree*, *Neural Network*, *K-Nearest Neighbour*, *Artificial Neural Network*, dan lain metode klasifikasi lain sebagainya. Metode yang digunakan pada penelitian ini menggunakan metode *Algoritma Naive Bayes* yang memanfaatkan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu ilmuwan yang memprediksi probabilitas di masa depan berdasarkan dengan pengalaman di masa sebelumnya atau masa lalu. Faktor-faktor yang mempengaruhi angka kelahiran sendiri yaitu faktor jenis kelamin laki-laki, jenis kelamin perempuan, dan pasangan usia subur. Angka kelahiran bayi sebagai indikator yang penting untuk mencerminkan keadaan derajat kesehatan di suatu masyarakat. Data yang digunakan dari tahun 1993 sampai dengan tahun 2022. Hasil analisis *Algoritma Naive Bayes* pada penelitian ini yaitu sebanyak 27 data diprediksi rendah dan 3 data diprediksi rendah pada tahun 2017, 2020, 2021 serta mendapatkan nilai akurasi (*accuracy*), *Recall* dan *Precision* sebesar 100%.

**Kata kunci :** *Data Mining*, Prediksi, *Algoritma Naive Bayes*

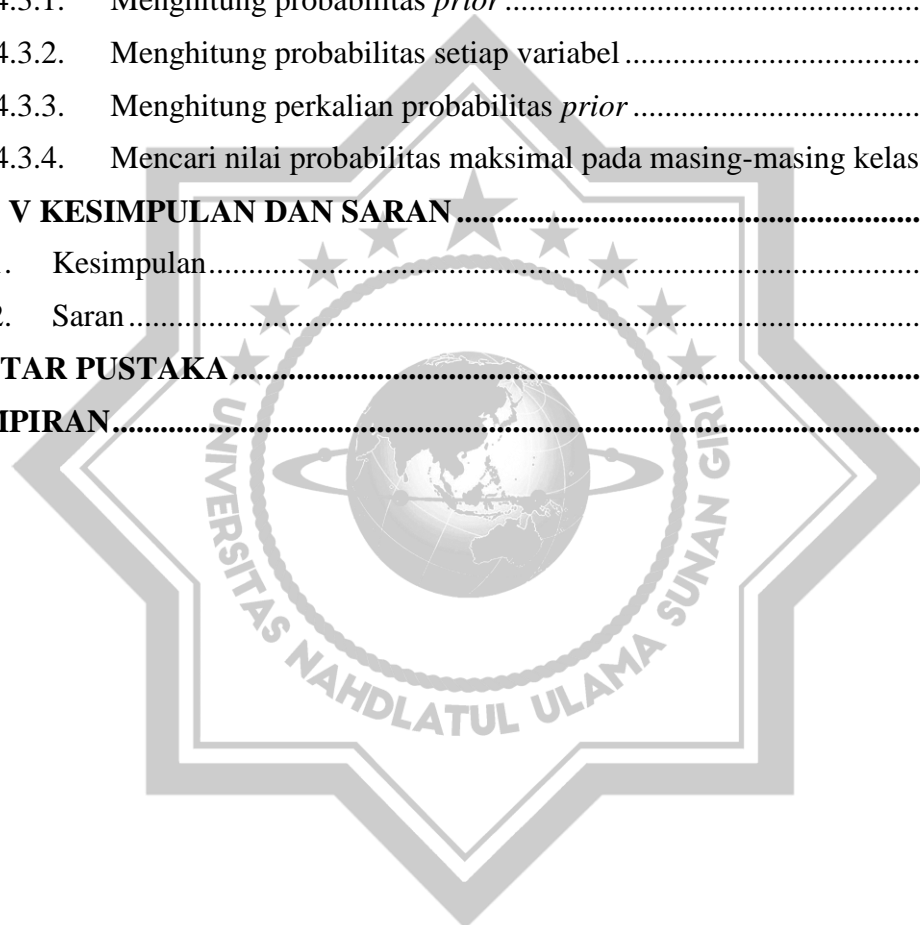
UNUGIRI



# DAFTAR ISI

	Halaman
<b>HALAMAN SAMPUL</b> .....	<b>i</b>
<b>HALAMAN SAMPUL DALAM</b> .....	<b>ii</b>
<b>PERNYATAAN</b> .....	<b>iii</b>
<b>HALAMAN PERSETUJUAN</b> .....	<b>iv</b>
<b>HALAMAN PENGESAHAN</b> .....	<b>v</b>
<b>MOTTO</b> .....	<b>vi</b>
<b>PERSEMBAHAN</b> .....	<b>vi</b>
<b>KATA PENGANTAR</b> .....	<b>vii</b>
<b>ABSTRACT</b> .....	<b>ix</b>
<b>ABSTRAK</b> .....	<b>x</b>
<b>DAFTAR ISI</b> .....	<b>xi</b>
<b>DAFTAR TABEL</b> .....	<b>xiii</b>
<b>DAFTAR BAGAN</b> .....	<b>xiv</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xv</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	6
1.3. Tujuan Penelitian.....	6
1.4. Manfaat Penelitian.....	6
<b>BAB II TINJAUAN PUSTAKA DAN DASAR TEORI</b> .....	<b>8</b>
2.1. Tinjauan Pustaka .....	8
2.2. Dasar Teori .....	14
2.2.1. Statistik Deskriptif .....	14
2.2.2. Ukuran Statistik Deskriptif .....	15
2.2.3. Data Mining .....	19
2.2.4. Prediksi.....	20
2.2.5. Algoritma Naive Bayes .....	21
2.2.6. Evaluasi Performansi Metode Klasifikasi.....	22
2.2.7. Angka Kelahiran Bayi.....	24
2.2.8. Ukuran Fertilitas.....	26
<b>BAB III METODE PENELITIAN</b> .....	<b>29</b>
3.1. Sumber Data .....	29

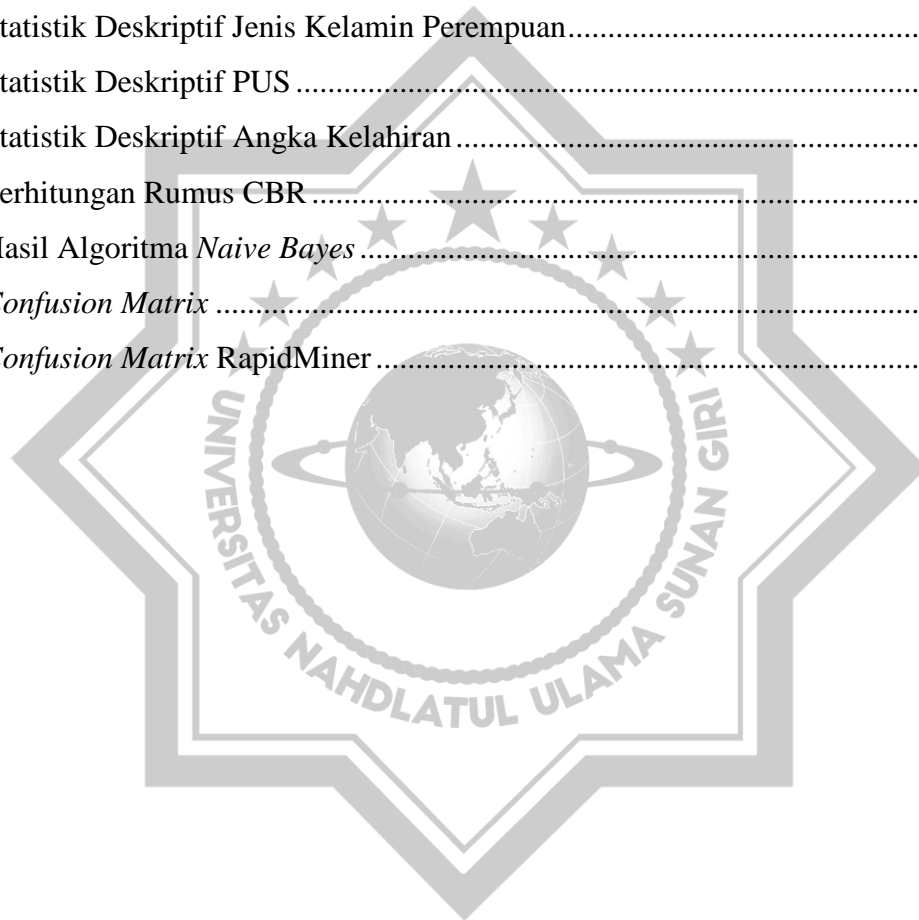
3.2. Variabel Penelitian .....	29
3.3. Teknik Analisis Data .....	30
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>33</b>
4.1. Statistik Deskriptif.....	33
4.2. Preprocessing Tingkat Kelahiran Bayi.....	35
4.3. Prediksi dan Tingkat Akurasi yang dihasilkan <i>Algoritma Naive Bayes</i> . 37	
4.3.1. Menghitung probabilitas <i>prior</i> .....	37
4.3.2. Menghitung probabilitas setiap variabel .....	39
4.3.3. Menghitung perkalian probabilitas <i>prior</i> .....	46
4.3.4. Mencari nilai probabilitas maksimal pada masing-masing kelas....	47
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>54</b>
5.1. Kesimpulan.....	54
5.2. Saran .....	55
<b>DAFTAR PUSTAKA .....</b>	<b>56</b>
<b>LAMPIRAN.....</b>	<b>60</b>



**UNUGIRI**

## DAFTAR TABEL

Tabel	Halaman
2.1 Penelitian Terdahulu .....	8
3.1 Variabel Penelitian .....	29
3.2 Struktur Data .....	30
4.1 Statistik Deskriptif Jenis Kelamin Laki-Laki .....	33
4.2 Statistik Deskriptif Jenis Kelamin Perempuan .....	34
4.3 Statistik Deskriptif PUS .....	34
4.4 Statistik Deskriptif Angka Kelahiran .....	34
4.5 Perhitungan Rumus CBR .....	35
4.6 Hasil Algoritma <i>Naive Bayes</i> .....	50
4.7 <i>Confusion Matrix</i> .....	51
4.8 <i>Confusion Matrix</i> RapidMiner .....	52



# UNUGIRI

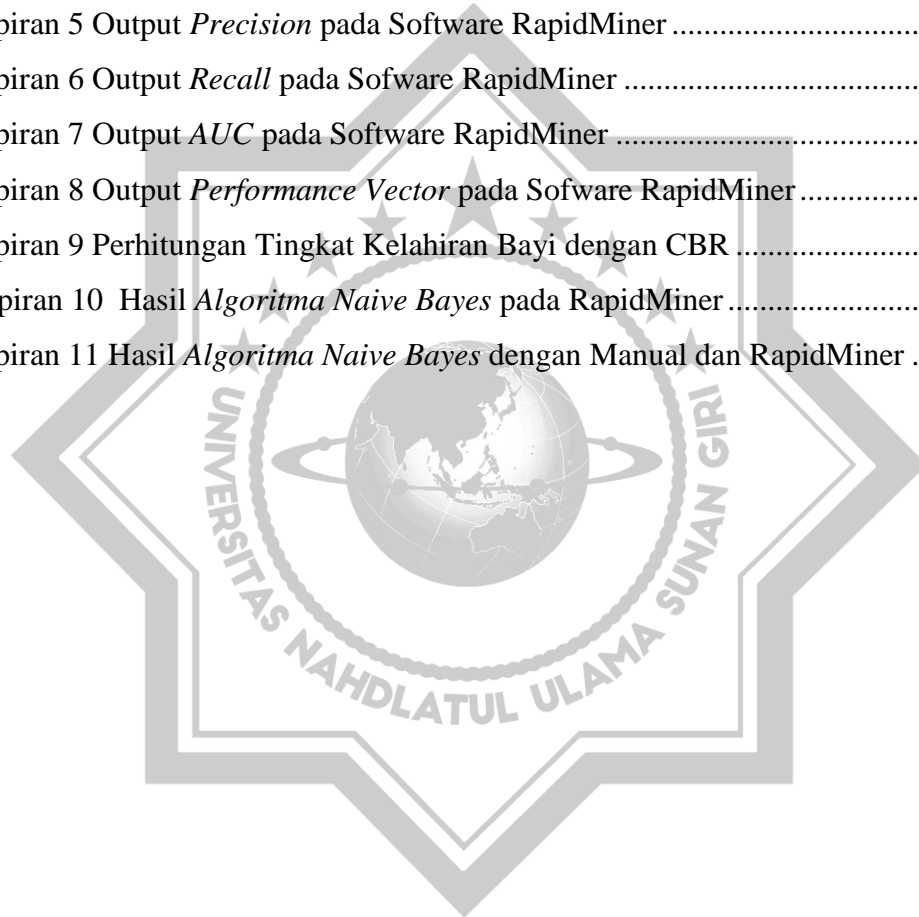
## DAFTAR BAGAN

Bagan	Halaman
3.1 Diagram Alir Pemodelan Naive Bayes .....	32



## DAFTAR LAMPIRAN

Lampiran	Halaman
Lampiran 1 Sebagian Data Penelitian.....	60
Lampiran 2 Output <i>Deskriptive Statistics</i> pada Software SPSS .....	60
Lampiran 3 Output Pemodelan <i>Naive Bayes</i> pada Software RapidMiner.....	61
Lampiran 4 Output <i>Accuracy</i> pada Software RapidMiner .....	61
Lampiran 5 Output <i>Precision</i> pada Software RapidMiner .....	61
Lampiran 6 Output <i>Recall</i> pada Software RapidMiner .....	62
Lampiran 7 Output <i>AUC</i> pada Software RapidMiner .....	62
Lampiran 8 Output <i>Performance Vector</i> pada Software RapidMiner .....	62
Lampiran 9 Perhitungan Tingkat Kelahiran Bayi dengan CBR .....	62
Lampiran 10 Hasil <i>Algoritma Naive Bayes</i> pada RapidMiner .....	62
Lampiran 11 Hasil <i>Algoritma Naive Bayes</i> dengan Manual dan RapidMiner .....	62



**UNUGIRI**