

**PERBANDINGAN METODE KLASIFIKASI REGRESI
LOGISTIK BINER DAN NAIVE BAYES PADA STATUS
PENGGUNA KB DI KOTA TEGAL TAHUN 2014**



SKRIPSI

Disusun Oleh :

NANCI RAJAGUKGUK

24010210141021

JURUSAN STATISTIKA

FAKULTAS SAINS DAN MATEMATIKA

UNIVERSITAS DIPONEGORO

SEMARANG

2015

**PERBANDINGAN METODE KLASIFIKASI REGRESI
LOGISTIK BINER DAN NAIVE BAYES PADA STATUS
PENGGUNA KB DI KOTA TEGAL TAHUN 2014**

Disusun Oleh :

NANCI RAJAGUKGUK

24010210141021

**Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Statistika pada Jurusan Statistika**

**JURUSAN STATISTIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG**

2015

HALAMAN PENGESAHAN I

Judul Skripsi : Perbandingan Metode Klasifikasi Regresi Logistik Biner
dan Naive Bayes pada Status Pengguna KB di Kota
Tegal Tahun 2014

Nama Mahasiswa : Nanci Rajagukguk

NIM : 24010210141021

Telah diujikan pada sidang Tugas Akhir tanggal 19 Maret 2015 dan dinyatakan
lulus pada tanggal 30 Maret 2015

Semarang, Maret 2015

Mengetahui,

Ketua Jurusan statistika

Fakultas Sains dan Matematika Undip



Dwi Sriyanti, M.Si
NIP. 195709141986032001

Panitia Penguji Ujian Tugas Akhir

Ketua



Hashbi Yasin, M.Si
NIP. 198212172006041003

HALAMAN PENGESAHAN II

Judul skripsi : Perbandingan Metode Klasifikasi Regresi Logistik Biner
dan Naive Bayes pada Status Pengguna KB di Kota
Tegal Tahun 2014

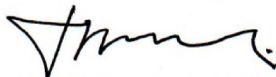
Nama Mahasiswa : Nanci Rajagukguk

NIM : 24010210141021

Telah diujikan pada sidang Tugas Akhir tanggal 19 Maret 2015

Semarang, Maret 2015

Pembimbing I



Dra. Dwi Ispriyanti, M.Si.
NIP.195709141986032001

Pembimbing II



Yuciana Wilandari, M.Si.
NIP.197005191998022001

KATA PENGANTAR

Puji syukur penulis ucapkan kepada Tuhan Yesus Kristus atas kasih dan anugerah-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul **“Perbandingan Metode Klasifikasi Regresi Logistik Biner dan Naive Bayes pada Status Pengguna KB di Kota Tegal Tahun 2014”**. Pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Ibu Dra. Dwi Ispriyanti, M.Si. selaku Ketua Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro Semarang sekaligus sebagai Dosen Pembimbing I.
2. Ibu Yuciana Wilandari, M.Si. sebagai pembimbing II yang telah memberikan bimbingan dan pengarahan kepada penulis.
3. Bapak dan Ibu dosen Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro yang telah memberikan ilmu yang bermanfaat
4. Semua pihak yang tidak dapat disebutkan satu per satu yang telah mendukung penulis menyelesaikan penulisan skripsi ini.

Penulis berharap Tugas Akhir ini dapat bermanfaat bagi seluruh civitas akademika di Universitas Diponegoro, khususnya Jurusan Statistika dan masyarakat pada umumnya.

Semarang, Maret 2015

Penulis

ABSTRAK

Indonesia merupakan negara yang termasuk memiliki jumlah penduduk terbanyak di dunia. Hal ini disebabkan karena negara Indonesia memiliki tingkat kelahiran yang begitu tinggi. Salah satu upaya penanganan agar pertumbuhan penduduk dapat dikendalikan yaitu dengan menggunakan program Keluarga Berencana (KB). Dalam penelitian ini, metode yang digunakan ialah Regresi Logistik Biner dan Naive Bayes. Untuk melakukan klasifikasi status pengguna KB pada kota Tegal Tahun 2014, variabel yang digunakan ialah umur istri, usia pertama menikah, jenis pekerjaan istri, jenis pekerjaan suami, status pendidikan istri, status pendidikan suami, dan jumlah anak. Perbandingan data training dan testing yang digunakan adalah 70:30. Berdasarkan hasil penelitian dengan menggunakan metode Regresi Logistik Biner diperoleh bahwa variabel prediktor yang signifikan yang berpengaruh terhadap status pengguna KB adalah umur istri, jenis pekerjaan istri, dan jumlah anak dengan ketepatan klasifikasi data testing yang dihasilkan sebesar 83,33%. Sedangkan dengan metode Naive Bayes diperoleh ketepatan klasifikasi sebesar 81,75%. Dari analisis tersebut dapat disimpulkan bahwa metode Regresi Logistik Biner lebih baik dibandingkan dengan Naive Bayes dalam mengklasifikasikan status pengguna KB di Kota Tegal tahun 2014.

Kata Kunci : Regresi Logistik Biner, Naive Bayes, Keluarga Berencana, Klasifikasi.

ABSTRACT

Indonesia is a country that includes having the highest population density in the world. It is because the Indonesian state has a birth rate is so high. One of the efforts to control that population growth can be controlled by using the Keluarga Berencana program. In this study, the method used is the Binary Logistic Regression and Naive Bayes. To perform classification KB User Status in Tegal 2014, the variable used is the wife's age, the age of first marriage, type of wife's job, type of husband's job, wife's education, husband's education, and number of children. The training data comparison testing is 70:30. Based on the research results using binary logistic regression showed that a significant predictor variables that affect the status of keluarga Berencana user are wife's age, type of wife's job, and number of children with a classification accuracy of testing data 83.33%. While with the Naive Bayes method obtained classification accuracy of 81.75%. From this analysis it can be concluded that the Binary Logistic Regression method is better than the Naive Bayes in classifying the status of KB users in Tegal 2014.

Keywords : Binary Logistic Regression, Naive Bayes, Keluarga Berencana, Classification.

DAFTAR ISI

| | Hal |
|---|-----|
| HALAMAN JUDUL..... | i |
| HALAMAN PENGESAHAN I | ii |
| HALAMAN PENGESAHAN II..... | iii |
| KATA PENGANTAR..... | iv |
| ABSTRAK..... | v |
| ABSTRACT..... | vi |
| DAFTAR ISI..... | vii |
| DAFTAR TABEL..... | x |
| DAFTAR GAMBAR..... | xi |
| DAFTAR LAMPIRAN..... | xii |
| BAB I PENDAHULUAN | |
| 1.1. Latar Belakang..... | 1 |
| 1.2. Rumusan Masalah..... | 3 |
| 1.3. Batasan Masalah..... | 3 |
| 1.4. Tujuan..... | 3 |
| BAB II TINJAUAN PUSTAKA | |
| 2.1. Kota Tegal..... | 4 |
| 2.1.1. Sejarah Kota Tegal..... | 4 |
| 2.1.2. Kondisi Geografis..... | 5 |
| 2.1.3. Pemerintahan dan Penduduk..... | 7 |
| 2.2. Keluarga Berencana..... | 7 |
| 2.2.1. Sejarah Keluarga Berencana di Indonesia..... | 7 |

| | |
|--|----|
| 2.2.2. Pengertian dan Tujuan Program KB..... | 9 |
| 2.2.3. Macam-Macam KB..... | 10 |
| 2.3. Konsep Klasifikasi dan Pengukurannya..... | 11 |
| 2.4. Regresi Logistik Biner..... | 13 |
| 2.4.1. Estimasi Parameter..... | 15 |
| 2.4.2. Pengujian Model..... | 18 |
| 2.5. Probabilitas dan Partisi..... | 20 |
| 2.6. Naive Bayes Klasifier..... | 22 |

BAB III METODOLOGI PENELITIAN

| | |
|----------------------------------|----|
| 3.1. Jenis dan Sumber Data..... | 35 |
| 3.2. Variabel Data..... | 35 |
| 3.3. Teknik Pengolahan Data..... | 36 |

BAB IV HASIL DAN PEMBAHASAN

| | |
|--|----|
| 1.1 Deskripsi Data..... | 39 |
| 1.2 Analisis Status Pengguna KB dengan Metode Regresi Logistik Biner..... | 41 |
| 4.2.1. Model Awal..... | 41 |
| 4.2.2. Uji Rasio Likelihood..... | 42 |
| 4.2.3 Uji Wald..... | 43 |
| 4.2.4. Uji Rasio Likelihood II..... | 45 |
| 4.2.5. Uji Wald II..... | 46 |
| 4.2.6 Uji Kesesuaian Model..... | 47 |
| 4.2.7 Model Akhir..... | 48 |
| 4.2.8 Estimasi dan Interpretasi Model Regresi Logistik Biner .. | 49 |

| | |
|---|----|
| 4.3. Ketepatan klasifikasi Menggunakan Regresi Logistik Biner.. | 51 |
| 4.4. Ketepatan Klasifikasi menggunakan Naive Bayes..... | 52 |
| 4.5. Perbandingan Ketepatan Klasifikasi..... | 53 |
| BAB V KESIMPULAN..... | 54 |
| DAFTAR PUSTAKA..... | 55 |
| LAMPIRAN..... | 56 |

DAFTAR TABEL

| | Hal |
|--|-----|
| Tabel 1. Matriks Konfusi Untuk Klasifikasi Dua Kelas..... | 12 |
| Tabel 2. Contoh Data Perhitungan Manual Naive Bayes..... | 25 |
| Tabel 3. Simulasi Hasil Prediksi dengan Metode Naive Bayes..... | 31 |
| Tabel 4. Status Pengguna dan Bukan Pengguna KB..... | 39 |
| Tabel 5. Variabel Umur Istri, Usia Kawin Pertama, Jumlah Anak..... | 39 |
| Tabel 6. Variabel Jenis Pekerjaan Istri, Jenis Pekerjaan Suami, Status Pendidikan Istri, Status Pendidikan Suami..... | 40 |
| Tabel 7. Nilai Estimasi Parameter..... | 42 |
| Tabel 8. Hasil Uji Wald..... | 44 |
| Tabel 9. Hasil Uji Wald II..... | 47 |
| Tabel 10. Nilai Odds Ratio..... | 49 |
| Tabel 11. Hasil Ketepatan Klasifikasi Regresi Logistik Biner..... | 51 |
| Tabel 12. Matriks Konfusi Naive Bayes..... | 52 |
| Tabel 13. Perbandingan Ketepatan Klasifikasi..... | 53 |

DAFTAR GAMBAR

| | Hal |
|---|-----|
| Gambar 1. Hasil Output Prediksi Naive Bayes Menggunakan Aplikasi Matlab 2009..... | 32 |
| Gambar 2. Diagram alir Pengolahan Data Regresi Logistik Biner dan Naive Bayes..... | 38 |

DAFTAR LAMPIRAN

| | Hal |
|--|-----|
| Lampiran 1. Data Status Pengguna KB di Kota Tegal Tahun 2014..... | 56 |
| Lampiran 2. Output Regresi Logistik Biner Semua Variabel..... | 57 |
| Lampiran 3. Output Regresi Logistik Biner Variabel yang Signifikan..... | 62 |
| Lampiran 4. Ketepatan Klasifikasi dengan Metode Regresi Logistik Biner.. | 67 |
| Lampiran 5. Output Metode Naive Bayes..... | 70 |
| Lampiran 6. Syntak Pengolahan Data Metode Naive Bayes..... | 72 |
| Lampiran 7. Tabel <i>Chi-square</i> | |

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara yang termasuk memiliki jumlah penduduk terbanyak di dunia. Hal ini disebabkan karena negara Indonesia memiliki tingkat kelahiran yang begitu tinggi sehingga terjadilah kepadatan penduduk. Kepadatan penduduk tersebut tentu saja menjadi suatu masalah bagi negara Indonesia yang perlu diperhatikan oleh pemerintah sehingga banyak upaya yang dipilih atau diprogramkan oleh pemerintah Indonesia untuk mengurangi kepadatan penduduk tersebut dengan cara melakukan program Keluarga Berencana atau dikenal dengan singkatan KB.

Badan Kependudukan Keluarga Berencana Nasional (BKKBN) merupakan lembaga pemerintah non kementerian yang berada di bawah dan bertanggungjawab kepada Presiden melalui Menteri yang bertanggungjawab dibidang Kesehatan. Sesuai Perpres 62 tahun 2010, BKKBN mempunyai tugas melaksanakan pemerintahan di bidang pengendalian penduduk dan penyelenggaraan keluarga berencana.

Menurut Entjang(1986) dalam Garis-garis Besar Haluan Negara (GBHN) sebagai Ketetapan Majelis Permusyawaratan Rakyat (MPR) No.IV/MPR/1987 disebutkan bahwa program KB bertujuan untuk meningkatkan kesejahteraan ibu dan anak dalam rangka mewujudkan keluarga bahagia yang menjadi dasar bagi terwujudnya masyarakat yang sejahtera dengan mengendalikan kelahiran sekaligus dalam rangka menjamin terkendalinya pertumbuhan penduduk Indonesia. Program

keluarga berencana dilaksanakan atas dasar sukarela serta tidak bertentangan dengan agama, kepercayaan dan moral Pancasila. Dengan demikian maka bimbingan, pendidikan serta pengarahan amat diperlukan agar masyarakat dengan kesadarannya sendiri dapat menghargai dan menerima pola keluarga kecil sebagai salah satu langkah utama untuk meningkatkan kesejahteraan hidupnya. Oleh karena itu pelaksanaan program keluarga berencana tidak hanya menyangkut masalah teknis medis semata-mata, melainkan meliputi berbagai segi penting lainnya dalam tata hidup dan kehidupan masyarakat.

Keberhasilan program KB terlihat dari jumlah pertumbuhan penduduk yang diharapkan. Ada banyak faktor yang mempengaruhi masyarakat dalam menggunakan KB. Dalam penelitian ini diasumsikan bahwa faktor-faktor yang mempengaruhi status menggunakan KB antara lain umur istri, umur pertama kawin istri, jenis pekerjaan istri, jenis pekerjaan suami, status pendidikan istri, status pendidikan suami dan jumlah anak. Pengguna KB diklasifikasikan dalam dua kategori yaitu pengguna dan tidak pengguna, sehingga ini dijadikan sebagai variabel respon (variabel tak bebas), dengan variabel responnya berupa data kualitatif dikotomi yaitu bernilai 1 untuk menyatakan keberadaan sebuah karakteristik dan bernilai 0 untuk menyatakan ketidakberadaan sebuah karakteristik. Salah satu metode statistika untuk mengklasifikasikan adalah Regresi Logistik Biner dan Naive Bayes. Dalam penelitian ini akan dibandingkan kedua metode tersebut. Pada penelitian ini akan diaplikasikan kedua metode tersebut pada keluarga berencana serta membandingkan keoptimalan kedua metode tersebut dalam mengklasifikasi data pengguna KB di Kota Tegal pada tahun 2014.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka dapat dirumuskan permasalahan yaitu bagaimana hasil perbandingan ketepatan pengklasifikasian pada metode Regresi Logistik Biner dan Naive Bayes sesuai dengan data Keluarga Berencana untuk Kota Tegal tahun 2014.

1.3 Batasan Masalah

Pembatasan masalah dalam penelitian ini adalah data Keluarga Berencana Kota Tegal yaitu sesuai dengan pendataan yang dilakukan oleh BKKBN Jawa Tengah tahun 2014. Pengolahan data tersebut menggunakan dua metode, yaitu metode Regresi Logistik Biner dan Naive Bayes .

1.4 Tujuan

Adapun tujuan dari penulisan tugas akhir ini adalah :

1. Memperoleh model dan ketepatan klasifikasi pengguna KB dengan menggunakan metode Regresi Logistik Biner
2. Memperoleh ketepatan klasifikasi dengan menggunakan metode Naive Bayes
3. Memperoleh perbandingan ketepatan klasifikasi Regresi Logistik Biner dan Naive Bayes pada data pengguna KB Kota Tegal tahun 2014.