



PERBANDINGAN REGRESI LOGISTIK DAN CHAID PADA FAKTOR RESIKO ANEMIA, DEFISIENSI BESI, DAN ANEMIA DEFISIENSI BESI WANITA USIA SUBUR

SHALY WANDA HAMZAH



PROGRAM STUDI STATISTIKA DAN SAINS DATA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024



(a) Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tesis dengan judul “Perbandingan Regresi Logistik dan CHAID pada Faktor Resiko Anemia, Defisiensi Besi, dan Anemia Defisiensi Besi Wanita Usia Subur” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir tesis ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2024

Shaly Wanda Hamzah
G1501202060



RINGKASAN

SHALY WANDA HAMZAH. Perbandingan Regresi Logistik dan CHAID pada Faktor Resiko Anemia, Defisiensi Besi, dan Anemia Defisiensi Besi Wanita Usia Subur. Dibimbing oleh MUHAMMAD NUR AIDI, I MADE SUMERTAJAYA, FITRAH ERNAWATI

Wanita usia subur (WUS) di Indonesia rentan terhadap anemia, defisiensi besi (DB), dan anemia defisiensi besi (ADB). Penelitian ini bertujuan untuk menentukan prevalensi kondisi tersebut pada WUS serta mengidentifikasi faktor-faktor yang memengaruhinya menggunakan analisis regresi logistik dan *Chi-squared Automatic Interaction Detector* (CHAID). Data yang digunakan berasal dari Riset Kesehatan Dasar 2013 yang disediakan oleh Badan Penelitian dan Pengembangan Kesehatan, Kementerian Kesehatan Indonesia, meliputi kadar hemoglobin, data demografi, dan sosial ekonomi, serta data ferritin (Fe) dan CRP dari sampel serum yang dikumpulkan pada tahun 2013, kemudian disimpan dan dianalisis pada tahun 2016.

Hasil penelitian menunjukkan bahwa prevalensi anemia pada WUS di Indonesia adalah 11%, DB 14%, dan ADB 9%. Peubah yang berpengaruh pada kondisi Anemia dengan menggunakan regresi logistik biner diantaranya adalah Indeks Massa Tubuh (IMT), konsumsi daging, penyakit malaria, dan ISPA. Pada saat menggunakan regresi logistic multinomial peubah yang berpengaruh pada anemia adalah IMT, status kawin, malaria, dan ISPA. Pada saat menggunakan metode CHAID peubah yang berpengaruh terhadap anemia meliputi penyakit malaria, jumlah konsumsi sayur mingguan, IMT, penyakit ISPA, jumlah konsumsi protein hewani (daging/ayam/ikan) harian, status kepemilikan rumah, status kehamilan, dan jumlah anggota keluarga.

Pada Kondisi DB, peubah yang berpengaruh dengan menggunakan regresi logistik biner diantaranya adalah IMT, pendidikan, besar keluarga, status kawin, dan penyakit diare. Pada saat menggunakan regresi logistic multinomial, peubah yang berpengaruh pada DB adalah IMT, pendidikan, besar keluarga, dan status kawin. Pada saat menggunakan metode CHAID peubah yang berpengaruh terhadap DB terdiri dari status kawin, status kehamilan, IMT, besar keluarga, penyakit diare, penyakit malaria, usia, jumlah konsumsi sayuran, pendidikan, kepemilikan rumah, penyakit ISPA, dan penyakit pneumonia.

Berdasarkan kondisi ADB, peubah yang signifikan pada saat menggunakan regresi logistic biner diantaranya adalah IMT, status rumah, besar keluarga, status kawin, daging, dan malaria. Pada saat menggunakan regresi logistic multinomial, peubah yang berpengaruh pada ADB adalah IMT, status rumah, besar keluarga, status kawin, daging, dan malaria. Pada saat menggunakan metode CHAID, peubah yang berpengaruh adalah IMT, status kawin, jumlah anggota keluarga, jumlah konsumsi daging, jumlah konsumsi buah mingguan, jumlah konsumsi sayuran, status kehamilan, penyakit ISPA, kepemilikan rumah, penyakit malaria, usia, penyakit diare, dan wilayah tempat tinggal.

Berdasarkan evaluasi model menunjukkan bahwa regresi logistik dan CHAID menghasilkan akurasi yang baik pada kondisi anemia 85%, DB 81%, dan ADB 88%. Namun, nilai sensitivitas kedua metode sangat tinggi pada 100% sementara spesifisitas sangat rendah yaitu 0% untuk kondisi DB dan ADB, yang menunjukkan



bahwa model dapat mengidentifikasi kondisi normal tetapi kurang efektif dalam mendeteksi kondisi anemia, DB, atau ADB. Tingginya data pada kategori normal menyebabkan model lebih cenderung mengidentifikasi kondisi normal dengan baik, namun kurang mampu mengidentifikasi kondisi anemia, DB, dan ADB. Ketika dianalisis dengan empat kategori, akurasi menurun menjadi sekitar 65% untuk kedua metode, dengan sensitivitas hampir 100% untuk kategori normal dan nol untuk kategori anemia, DB, dan ADB. Hasil dari model yang didapatkan maka disimpulkan model dengan metode regresi loistik maupun CHAID memiliki akurasi tinggi yang mampu mendeteksi kondisi normal namun belum sensitif untuk mendeteksi kondisi anemia, defisiensi besi maupun anemia defisiensi besi. Hal ini terjadi disebabkan oleh tidak seimbangnya data normal dengan data anemia, DB maupun ADB.

Kata kunci: Anemia, CHAID, Defisiensi Besi, Regresi Logistik, Wanita Usia Subur



SUMMARY

SHALY WANDA HAMZAH. Comparison of Logistic Regression and CHAID on Risk Factors for Anemia, Iron Deficiency, and Iron Deficiency Anemia in Women Reproductive Age. Supervised by MUHAMMAD NUR AIDI, I MADE SUMERTAJAYA, FITRAH ERNAWATI

Women of reproductive age (WRA) in Indonesia are vulnerable to anemia, iron deficiency (ID), and iron deficiency anemia (IDA). This study aims to determine the prevalence of these conditions among WRA and identify the influencing factors using logistic regression analysis and *Chi-squared Automatic Interaction Detector* (CHAID). The data used is from the 2013 Basic Health Research provided by the Health Research and Development Agency, Ministry of Health Indonesia, which includes hemoglobin levels, demographic and socio-economic data, as well as ferritin (Fe) and CRP data from serum samples collected in 2013 and subsequently stored and analyzed in 2016.

The study results show that the prevalence of anemia among WRA in Indonesia is 11%, ID is 14%, and IDA is 9%. Variables influencing anemia using binary logistic regression include Body Mass Index (BMI), meat consumption, malaria, and acute respiratory infections (ARI). When using multinomial logistic regression, the influencing variables for anemia are BMI, marital status, malaria, and ARI. Using the CHAID method, the variables influencing anemia include malaria, weekly vegetable consumption, BMI, ARI, daily consumption of animal protein (meat/chicken/fish), homeownership status, pregnancy status, and family size.

For ID, variables influencing the condition identified using binary logistic regression include BMI, education, family size, marital status, and diarrhea. In multinomial logistic regression, the influencing variables for ID are BMI, education, family size, and marital status. Using the CHAID method, the variables influencing ID consist of marital status, pregnancy status, BMI, family size, diarrhea, malaria, age, vegetable consumption, education, homeownership status, ARI, and pneumonia.

For IDA, significant variables identified using binary logistic regression include BMI, homeownership status, family size, marital status, meat consumption, and malaria. In multinomial logistic regression, the influencing variables for IDA are BMI, homeownership status, family size, marital status, meat consumption, and malaria. Using the CHAID method, the influencing variables for IDA include BMI, marital status, family size, meat consumption, weekly fruit consumption, vegetable consumption, pregnancy status, ARI, homeownership status, malaria, age, diarrhea, and region of residence.

Model evaluation shows that logistic regression and CHAID yield good accuracy for anemia 85%, ID 81%, and IDA 88%. However, the sensitivity values for both methods are very high at 100%, while specificity is very low at 0% for ID and IDA, indicating that the models can identify normal conditions well but are less effective in detecting anemia, ID, or IDA. The high data in the normal category causes the models to better identify normal conditions but less effectively detect anemia, ID, and IDA. When analyzed with four categories, accuracy decreases to around 65% for both methods, with sensitivity nearly 100% for the normal category

and zero for anemia, ID, and IDA categories. The results of the models indicate that both logistic regression and CHAID methods have high accuracy in detecting normal conditions but are not yet sensitive enough to detect anemia, iron deficiency, or iron deficiency anemia. This occurs due to an imbalance between normal data and anemia, DB and ADB data.

Keywords: Anemia, CHAID, Iron Deficiency, Logistic Regression, Reproductive-Age Women.

@Hak cipta milik IPB University

IPB University





Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2024
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.



PERBANDINGAN REGRESI LOGISTIK DAN CHAID PADA FAKTOR RESIKO ANEMIA, DEFISIENSI BESI, DAN ANEMIA DEFISIENSI BESI WANITA USIA SUBUR

SHALY WANDA HAMZAH

Tesis
sebagai salah satu syarat untuk memperoleh gelar
Magister pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI STATISTIKA DAN SAINS DATA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Tesis:
Dr. Yenni Angraini, S.Si., M.Si.





Judul Tesis : Perbandingan Regresi Logistik dan CHAID pada Faktor Resiko
Anemia, Defisiensi Besi, dan Anemia Defisiensi Besi Wanita
Usia Subur

Nama : Shaly Wanda Hamzah

NIM : G1501202060

Disetujui oleh

Pembimbing 1:

Prof. Dr. Ir. Muhammad Nur Aidi, MS.

Pembimbing 2:

Dr. I Made Sumertajaya, M.Si.

Pembimbing 2:

Dr. Fitrah Ernawati, M.Sc



Digitally signed by:
Muhammad Nur Aidi

Date: 9 Agu 2024 14:44:18 WIB
Verify at dsign.ipb.ac.id



Diketahui oleh

Ketua Program Studi:

Dr. Agus M. Soleh

NIP. 197503151999031004

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam:

Dr. Berry Juliandi, S.Si., M.Si.

NIP. 197807232007011001



Digitally signed

dsign.ipb.ac.id



Digitally signed

dsign.ipb.ac.id

Tanggal Ujian: 16 Juli 2024

Tanggal Lulus:



(a) Hak cipta milik IPB University

IPB University



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan September 2022 sampai bulan Juni 2024 ini ialah gizi, dengan judul “Perbandingan Regresi Logistik dan CHAID pada Faktor Resiko Anemia, Defisiensi Besi, dan Anemia Defisiensi Besi Wanita Usia Subur”.

Terima kasih penulis ucapkan kepada para pembimbing, Prof. Dr. Ir. Muhammad Nur Aidi, M.S., Dr. I Made Sumertajaya, M.Si., dan Dr. Fitrah Ernawati, M.Sc. yang telah membimbing dan banyak memberi saran pada penelitian ini. Ucapan terima kasih juga disampaikan kepada penguji luar komisi pembimbing Dr. Yenni Angraini, S.Si., M.Si yang telah banyak memberikan saran pada penulisan karya ilmiah ini. Penghargaan penulis sampaikan kepada Badan Riset dan Inovasi Nasional yang telah menyediakan data penelitian, serta membimbing dalam proses penulisan. Ungkapan terima kasih juga disampaikan kepada Mama, Papa, Abi, dan seluruh keluarga dan sahabat yang telah memberikan dukungan, doa, dan kasih sayang.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2024

Shaly Wanda Hamzah



(a) Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University



DAFTAR ISI

DAFTAR TABEL	x
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	x
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
II TINJAUAN PUSTAKA	3
2.1 Analisis Regresi Logistik	3
2.2 <i>Chi-square Automatic Interaction Detection (CHAID)</i>	4
2.3 Evaluasi Kebaikan Model	5
2.4 Anemia dan Defisiensi Besi	6
III METODE	8
3.1 Data Penelitian	8
3.2 Peubah Penelitian	8
3.3 Tahapan Analisis Data	9
IV HASIL DAN PEMBAHASAN	11
4.1 Analisis Deskriptif	11
4.2 Analisis Regresi Logistik	15
4.3 CHAID	19
4.4 Evaluasi Model	27
V SIMPULAN	29
DAFTAR PUSTAKA	30
LAMPIRAN	34
RIWAYAT HIDUP	52



DAFTAR TABEL

1	Matriks Konfusi	6
2	Status WUS berdasarkan nilai Hb dan Fe	7
3	Peubah dependen penelitian	8
4	Peubah penelitian	8
5	Tabulasi silang status anemia menurut faktor risikonya	11
6	Tabulasi silang status DB menurut faktor risikonya	12
7	Tabulasi silang status ADB menurut faktor risikonya	13
8	Tabulasi silang status anemia, DB, dan, ADB menurut faktor risikonya	14
9	Hasil Analisis regresi logistik biner peubah dependen Y_1 , Y_2 , dan Y_3 terhadap peubah independen X_1 sd X_{17}	15
10	Hasil Analisis regresi logistik multinomial Y_4 terhadap peubah independen X_1 sd X_{17}	17
11	Pohon klasifikasi CHAID berdasarkan kategori anemia	19
12	Pohon klasifikasi CHAID berdasarkan kategori DB	21
13	Pohon klasifikasi CHAID berdasarkan kategori ADB	24
14	Pohon klasifikasi CHAID berdasarkan kategori anemia, DB dan ADB	25
15	Perbandingan kinerja metode regresi logistik dan CHAID	27

DAFTAR GAMBAR

1	Hubungan Anemia, DB dan ADB (WHO,2015)	7
2	WUS berdasarkan kadar Hb dan Fe	11

DAFTAR LAMPIRAN

1	Lampiran 1 Tabulasi silang Y_1 status Anemia – Normal	35
2	Lampiran 2 Tabulasi silang Y_2 status DB – Normal	37
3	Lampiran 3 Tabulasi silang Y_3 status ADB - Normal	39
4	Lampiran 4 Tabulasi silang Y_4 status Normal, Anemia, DB dan ADB	41
5	Lampiran 5 Regresi Logistik Biner	44
6	Lampiran 6 Regresi Logistik Multinomial	46
7	Lampiran 7 CHAID peubah dependen Y_1	48
8	Lampiran 8 CHAID peubah dependen Y_2	49
9	Lampiran 9 CHAID peubah dependen Y_3	50
10	Lampiran 10 CHAID peubah dependen Y_4	51