

ANALISIS KLASIFIKASI KELAHIRAN CAESAR MENGUNAKAN ALGORITMA NAÏVE BAYES

Ginanjar Abdurrahman¹, Johan Taruna Wijaya²
Dosen Universitas Muhammadiyah Jember¹,
Mahasiswa Universitas Muhammadiyah Jember²
Jalan Karimata No.49, Jember

Pos-el : abdurrahmanganjar@gmail.com¹, johantaruna@gmail.com²

Abstract

The maternal mortality rate (MMR) in Indonesia is still high. AKI is the number of maternal deaths during pregnancy, childbirth and childbirth caused by pregnancy, childbirth and childbirth or management and is not caused by accident or falls in every 100,000 live births. Caesarean birth is the last alternative in the act of labor. This is because of the high risk factors, both risks for both mother and baby. Despite the high risk, Caesar's birth rate has increased significantly, especially in Indonesia. The World Health Organization (WHO) sets a standard for Caesarean delivery in a country of around 5-15 percent per 1000 births in the world. Machine learning learns how machines or computers learn from experience or how to program machines to learn. Machine learning requires data to learn or in other words termed learn from data. There are several machine learning applications that have been developed. One is WEKA machine learning (Waikato Environment for Knowledge Analysis). Weka is a data mining software that is equipped with a standard machine learning algorithm, including preprocessing, classification, clustering, regression, association, and visualization algorithms. In this study the classification of Caesar's birth will be carried out using the Naïve Bayes algorithm to determine whether the birth will be given an act of Caesarean section or a normal delivery (No Caesar). The dataset used consists of 80 data of pregnant women with 5 attributes, namely: age, number of medical personnel, delivery time, blood pressure, liver problems. The results of this study are expected to help the health world especially to determine the birth process need to be given Caesarean section or not by providing a classification model to make medical predictions. The results of this study are expected to be able to support the world of health, especially in the case of giving birth by providing a classification model that can be used to make predictions and other medical decision making.

Keywords: Birth, Caesar, Classification, Naive Bayes, Weka

I. Pendahuluan

Angka kematian Ibu (AKI) di Indonesia masih tinggi. AKI merupakan jumlah kematian ibu selama masa kehamilan, persalinan dan nifas yang

disebabkan oleh kehamilan, persalinan dan nifas atau pengelolaannya dan tidak disebabkan karena kecelakaan atau terjatuh di setiap 100.000 kelahiran hidup (Kemenkes, 2017).

Adapun upaya yang dapat dilakukan untuk menurunkan AKI diantaranya (Kemenkes, 2017):

1. Pelayanan kesehatan ibu hamil
2. Pertolongan persalinan oleh tenaga kesehatan terlatih di fasilitas pelayanan kesehatan
3. Perawatan pasca persalinan bagi ibu dan bayi
4. Perawatan khusus dan rujukan jika terjadi komplikasi
5. Kemudahan mendapatkan cuti hamil dan melahirkan
6. Pelayanan keluarga berencana

Kelahiran secara Caesar merupakan alternatif terakhir dalam tindakan persalinan. Hal ini dikarenakan faktor resiko yang tinggi, baik resiko bagi ibu maupun bayi (Onggang dalam (Anggorowati & Sudiharjani, 2010). Walaupun resikonya tinggi, angka kelahiran secara Caesar mengalami peningkatan secara signifikan, khususnya di Indonesia (Andayasari et al., 2015). World Health Organization (WHO) menetapkan standar persalinan operasi Caesar di suatu negara sekitar 5-15 persen per 1000 kelahiran di dunia (Sihombing, Saptarini, & Putri, 2010). Berdasarkan data WHO (Dalam (Andayasari et al., 2015)), pada tahun 2004-2008 di tiga benua (Amerika latin, Afrika dan Asia)

angka kelahiran Caesar terendah adalah di Angola (2,3%) dan tertinggi di Cina (46,2%). Data kelahiran Caesar di Indonesia meningkat tajam, khususnya di kota-kota besar. Berdasarkan data Riskedas (2010) dalam Andayasari (2015:106) kelahiran Caesar sebanyak 15,3%. Angka terendah di Sulawesi Tenggara (5,5%) dan tertinggi di DKI Jakarta (27,2%)

Machine learning mempelajari cara mesin atau komputer belajar dari pengalaman atau cara memprogram mesin untuk dapat belajar. Machine learning memerlukan data untuk belajar atau dengan kata lain diistilahkan dengan learn from data (Alpaydin dalam (Lukman & Marwana, 2015)). Terdapat beberapa aplikasi machine learning yang telah dikembangkan. Salah satu adalah *machine learning* WEKA (Waikato Environment for Knowledge Analysis). Weka merupakan software untuk data mining yang dilengkapi dengan algoritma standar machine learning, diantaranya algoritma preprocessing, klasifikasi, clustering, regresi, asosiasi, dan visualisasi (Lukman & Marwana, 2015)

Pada penelitian ini akan dilakukan klasifikasi kelahiran Caesar menggunakan algoritma Naïve Bayes untuk menentukan apakah kelahiran akan diberikan tindakan operasi Caesar atau melahirkan secara

normal (Tidak Caesar). Dataset yang digunakan terdiri dari 80 data ibu hamil dengan 5 atribut, yakni: umur, jumlah tenaga medis, waktu melahirkan, tekanan darah, masalah hati. Hasil penelitian ini diharapkan dapat membantu dunia kesehatan khususnya untuk menentukan proses kelahiran perlu diberi tindakan operasi Caesar ataukah tidak dengan menyediakan model klasifikasi untuk melakukan prediksi secara medis

II. Metodologi Penelitian

Identifikasi Masalah

Dalam langkah identifikasi masalah, dilakukan untuk mengenali masalah yang ada dalam penelitian berdasarkan latar belakang penelitian sehingga dapat ditawarkan solusi yang dapat digunakan untuk menyelesaikan masalah tersebut.

Studi Literatur

Studi literature merupakan langkah yang dilakukan untuk mempelajari referensi berupa jurnal penelitian, paper, buku-buku referensi yang lain terkait dengan penelitian untuk melengkapi pengetahuan awal, guna memahami teori yang dapat digunakan untuk menunjang penelitian.

Dataset

Dataset yang digunakan adalah data wanita hamil yang nantinya akan diklasifikasikan apakah akan melahirkan dengan tindakan Caesar ataukah secara normal. Data ini diperoleh dari [https://archive.ics.uci.edu/ml/datasets/Caesarian+Section+ Classification+ Dataset#](https://archive.ics.uci.edu/ml/datasets/Caesarian+Section+Classification+Dataset#). Dataset ini mempunyai 80 record dengan 6 variabel yaitu: age(umur), delivery number (jumlah tenaga medis yang terlibat), delivery time (waktu melahirkan), blood of pressure (tekanan darah), heart problem (masalah jantung), dan Caesar.

Data Preprocessing

Sebelum data digunakan untuk tahapan klasifikasi, dataset dilakukan data prerprocessing yang nantinya meliputi : pembersihan data yang tidak lengkap (*missing value*), transformasi data, dan konversi data agar dapat digunakan pada tahap klasifikasi.

Klasifikasi: Algoritma Naïve Bayes

Klasifikasi pada penelitian ini menggunakan bantuan program aplikasi WEKA 3.6, merupakan perangkat lunak open source yang telah dikenal dan umum digunakan di bidang *data mining* dan *machine learning*. Pada WEKA 3.6 terdapat banyak tersedia *library – library* algoritma, baik untuk klasifikasi, *clustering*, *association rule*, bahkan sampai

library untuk praproses data dan *feature selection*.

Klasifikasi nantinya akan menerapkan algoritma Naïve Bayes, dengan skenario uji dataset menggunakan pembagian persentase (*percentage split*), yaitu membagi jumlah keseluruhan data menjadi data *training* dan data *testing*. Adapun *percentage split* yang digunakan adalah: 5%, 15%, 25%, 35%, 45%, 55%, 65%, 75%, 85%, dan 95%. *Percentage split* 5% artinya 5% data digunakan sebagai data *training*, dan 95% data digunakan sebagai data *testing*, demikian juga untuk *percentage split* yang lainnya. Hal ini dilakukan agar hasil yang diperoleh obyektif.

Interpretasi dan Evaluasi

Tahap Interpretasi dan Evaluasi merupakan tahap untuk menggali informasi berupa pengetahuan dari hasil klasifikasi yang selanjutnya dievaluasi apakah data yang diklasifikasikan sudah benar dengan menggunakan *TP Rate*, *FP Rate*, *Precision* dan *Recall*

III. Hasil Dan Pembahasan

Langkah awal yang dilakukan pada penelitian ini adalah *preprocessing* data. Dataset yang digunakan adalah data ibu hamil yang nantinya akan dilakukan klasifikasi apakah melahirkan secara Caesar atau melahirkan secara normal.

Preprocessing data

Pada tahap ini, dataset yang digunakan dilakukan pembersihan data yang tidak lengkap (*missing value*), transformasi data, dan konversi data agar dapat digunakan pada tahap klasifikasi. Pada dataset ibu hamil ini, tidak terdapat *missing value*, outlier, maupun noise data. Dengan demikian, proses pembersihan data tidak perlu dilakukan. Selanjutnya, dilakukan transformasi dan konversi dataset.

Pada tahap transformasi dan konversi, format atau susunan data diubah menjadi format delimited, dengan pemisahannya adalah *coma* (,) dan menyimpan file ke dalam ekstensi *coma separated value* (*csv*). Tahap selanjutnya adalah konversi bentuk file ke dalam ekstensi (**.arff*) sehingga file dapat dipanggil dengan perangkat lunak WEKA.

Pada tahap ini, tahap *preprocessing* data selesai dilakukan, sehingga data dapat digunakan untuk tahap selanjutnya, yakni klasifikasi.

Klasifikasi

Pada tahap ini, digunakan aplikasi WEKA 3.9 sebagai perangkat untuk membantu proses klasifikasi. Algoritma klasifikasi yang digunakan pada penelitian ini adalah algoritma Naïve Bayes. Dataset

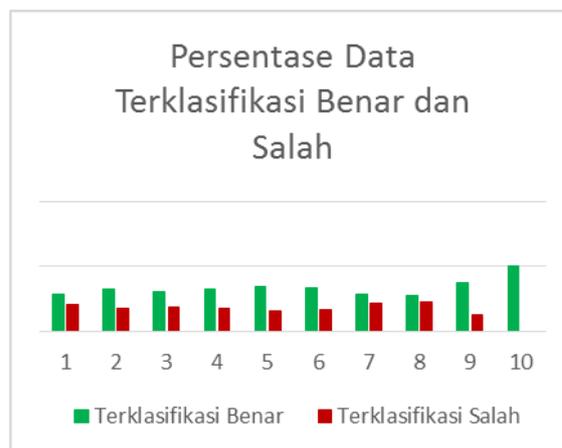
dibagi menjadi data *training* dan data *testing* menggunakan *percentage split*, sehingga diperoleh *percentage split* sebagai berikut: 5%, 15%, 25%, 35%, 45%, 55%, 65%, 75%, 85%, dan 95%. Hasil uji coba pada dataset dengan pembagian data menggunakan *percentage split* terlihat pada Tabel 1.

Dari Tabel 1 dapat dilihat rincian persentase data terklasifikasi benar dan persentase data terklasifikasi salah. Nilai rata-rata data terklasifikasi benar adalah 67,16%, sedangkan rata-rata data terklasifikasi salah adalah 32,84%. Adapun nilai tertinggi dan nilai terendah dari persentase data terklasifikasi benar masing-masing adalah 100% dan 55%. Sedangkan nilai tertinggi dan terendah dari persentase data terklasifikasi salah masing-masing adalah 45% dan 0%. Secara rata-rata, performa algoritma klasifikasi dengan algoritma Naïve Bayes masih kurang baik.

Tabel 1 Hasil uji coba dataset

| No | Percentage Split | Jumlah Data Training | Jumlah Data Testing | Persentase Data Terklasifikasi Benar | Persentase Data Terklasifikasi Salah |
|----|------------------|----------------------|---------------------|--------------------------------------|--------------------------------------|
| 1 | 5% | 4 | 76 | 57,89% | 42,1% |
| 2 | 15% | 12 | 68 | 64,71% | 35,29% |
| 3 | 25% | 20 | 60 | 61,67% | 38,33% |
| 4 | 35% | 28 | 52 | 65,38% | 34,62% |
| 5 | 45% | 36 | 44 | 68,18% | 31,82% |
| 6 | 55% | 44 | 36 | 66,67% | 33,33% |
| 7 | 65% | 52 | 28 | 57,14% | 42,86% |
| 8 | 75% | 60 | 20 | 55% | 45% |
| 9 | 85% | 68 | 12 | 75% | 25% |
| 10 | 95% | 76 | 4 | 100% | 0% |

Di samping hasil uji dari persentase data terklasifikasi benar dan persentase data terklasifikasi salah, kinerja algoritma Naïve Bayes untuk mengklasifikasikan kelahiran ibu hamil secara normal ataukah Caesar juga dilihat dari nilai *precision* dan *recall*. Pada penelitian ini data positif adalah data kelas melahirkan secara normal, sedangkan data negatif adalah data kelas melahirkan dengan Caesar. Nilai *precision* dan *recall* dapat dilihat pada Tabel 3

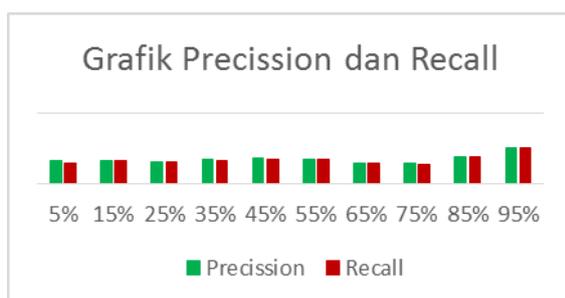


Tabel 2 Nilai TP Rate, FP Rate, Precision, dan Recall

| No. | Percentage Split | Jumlah Data Training | Jumlah Data Testing | TP Rate | FP Rate | Precision | Recall |
|-----|------------------|----------------------|---------------------|---------|---------|-----------|--------|
| 1 | 5% | 4 | 76 | 0,579 | 0,366 | 0,634 | 0,579 |
| 2 | 15% | 12 | 68 | 0,647 | 0,43 | 0,652 | 0,647 |
| 3 | 25% | 20 | 60 | 0,617 | 0,422 | 0,611 | 0,617 |
| 4 | 35% | 28 | 52 | 0,654 | 0,333 | 0,684 | 0,654 |
| 5 | 45% | 36 | 44 | 0,682 | 0,316 | 0,704 | 0,682 |
| 6 | 55% | 44 | 36 | 0,667 | 0,368 | 0,667 | 0,667 |
| 7 | 65% | 52 | 28 | 0,571 | 0,453 | 0,568 | 0,571 |
| 8 | 75% | 60 | 20 | 0,55 | 0,45 | 0,567 | 0,55 |
| 9 | 85% | 68 | 12 | 0,75 | 0,293 | 0,75 | 0,75 |
| 10 | 95% | 76 | 4 | 1 | 0 | 1 | 1 |

Berdasarkan Tabel 2, dapat dilihat nilai TP Rate, FP Rate, Precision dan Recall untuk masing-masing percentage

split. Sehingga diperoleh nilai rata-rata *TP Rate* sebesar 0,67, nilai rata-rata *FP Rate* sebesar 0,34, nilai rata-rata *Precision* sebesar 0,68, dan nilai rata-rata *Recall* sebesar 0,67. Berdasarkan nilai rata-rata *Precision* dan *Recall* dapat disimpulkan bahwa tingkat keefektifan klasifikasi dengan *Naïve Bayes* berada pada tingkat cukup baik. Nilai *precision* dan *recall* tertinggi berada pada hasil uji dataset dengan *percentage split* 95%, yakni masing-masing bernilai 1.



IV. Kesimpulan

Dari hasil penelitian yang telah diperoleh, dapat disimpulkan bahwa klasifikasi kelahiran menggunakan algoritma *Naïve Bayes* masih menghasilkan nilai yang kurang baik. Hal ini ditunjukkan dari rata-rata terklasifikasi benar bernilai adalah 67,16%, sedangkan rata rata data terklasifikasi salah adalah 32,84%. Adapun nilai tertinggi dan nilai terendah dari persentase data terklasifikasi benar masing-masing adalah 100% dan 55%. Sedangkan nilai tertinggi dan

terendah dari persentase data terklasifikasi salah masing-masing adalah 45% dan 0%. Nilai *precision* dan *recall* tertinggi berada pada hasil uji dataset dengan *percentage split* 95%, yakni masing-masing bernilai 1.

Berdasarkan hasil penelitian yang telah dilakukan, untuk penelitian selanjutnya disarankan untuk menerapkan algoritma klasifikasi lain sebagai pembandingan sehingga dapat meningkatkan performa klasifikasi.

DAFTAR RUJUKAN

- Andayasari, L., Muljati, S., Sihombing, M., Arlinda, D., Opitasari, C., Fajar, D., & Widiyanto, M. (2015). Proporsi Seksio Sesarea dan Faktor yang Berhubungan dengan Seksio Sesarea di Jakarta THE PROPORTION OF CAESAREAN SECTION AND ASSOCIATED FACTORS IN HOSPITAL OF JAKARTA. *Buletin Penelitian Kesehatan*, 43(2), 105–116.
<https://doi.org/http://dx.doi.org/10.22435/bpk.v43i2.4144.105-116>
- Anggorowati, & Sudiharjani, N. (2010). Mobilisasi Dini dan Penyembuhan Luka Operasi Pada Ibu Post Sectio Caesarea (SC) di Ruang Dahlia Rumah Sakit Umum Daerah Kota Salatiga. *Prosiding Seminar Nasional Dan Internasional Universitas Muhammadiyah Semarang*, 30–35.
- Kemenkes. (2017). *Profil Kesehatan Indonesia Tahun 2016*.
<https://doi.org/10.1111/evo.12990>
- Lukman, A., & Marwana. (2015). Machine

Learning Multi Klasifikasi.
Konferensi Nasional Ilmu Komputer (KONIK) 2014, (December 2014).

Sihombing, N., Saptarini, I., & Putri, D. S. K. (2010). Determinan Persalinan Sectio Caesarea Di Indonesia (Analisis Lanjut Data Riskesdas 2013). *Jurnal Kesehatan Reproduksi*, 8(1), 63–75.
<https://doi.org/10.22435/kespro.v8i1.6641.63-75>