

## KLASIFIKASI DATA BERAT BAYI LAHIR MENGGUNAKAN *PROBABILISTIC NEURAL NETWORK DAN REGRESI LOGISTIK* (Studi Kasus di Rumah Sakit Islam Sultan Agung Semarang Tahun 2014)

Erfan Sofha<sup>1</sup>, Hasbi Yasin<sup>2</sup>, Rita Rahmawati<sup>3</sup>

<sup>1</sup>Mahasiswa Jurusan Statistika FSM Universitas Diponegoro

<sup>2,3</sup>Staff Pengajar Jurusan Statistika FSM Universitas Diponegoro

[erfansofha072010@gmail.com](mailto:erfansofha072010@gmail.com), [hasbiyasin17@gmail.com](mailto:hasbiyasin17@gmail.com), [ritarahmawati@gmail.com](mailto:ritarahmawati@gmail.com)

### ABSTRACT

Birth Weight Infant (BWI) is the baby's weight weighed in an hour after being born. Factors that may influence the BWI such as maternal age, length of gestation, body weight, height, blood pressure, hemoglobin and parity. One possibility of BWI is Low Birth Weight Infant (LBWI) ( $BWI < 2500$  gram). LBWI is one of the causes of infant mortality. This study use the Probabilistic Neural Network (PNN) and Logistic Regression to classify the birth weight of infant in RSI Sultan Agung Semarang along the year of 2014. This study's aims are to know the factors that affect the BWI by using logistic regression and finally finding the best method between PNN and logistic regression methods in classifying the BWI data. As a result, gestation, body weight and hemoglobin are the factors that affect the BWI in RSI Sultan Agung Semarang. The accuracy of PNN classification method on training data is 100%, which is better than the logistic regression method giving only about 88,2%, while the testing data has the same great accuracy at 86,67%.

**Keywords:** BWI, LBWI, PNN, Logistic Regression, Classification

### 1. PENDAHULUAN

Salah satu indikator untuk mengetahui derajat kesehatan masyarakat adalah Angka Kematian Neonatal (AKN). AKN adalah jumlah bayi yang meninggal satu bulan pertama setelah kelahiran (0-28 hari) yang dinyatakan dalam 1.000 kelahiran hidup pada tahun yang sama. AKN di Indonesia periode 5 tahun terakhir sejak 2007 mengalami stagnansi. Berdasarkan laporan SDKI (Survei Demografi dan Kesehatan Indonesia) 2007 dan 2012 AKN diestimasikan sebesar 19 per 1.000 kelahiran hidup<sup>[5]</sup>. Berdasarkan perkiraan organisasi kesehatan dunia *World Health Organization* (WHO) hampir semua (98%) dari lima juta kematian neonatal terjadi di negara berkembang. Lebih dari dua pertiga kematian itu terjadi pada periode neonatal dini (0-7 hari), yang umumnya dikarenakan Berat Bayi Lahir (BBL) kurang dari 2.500 gram<sup>[2]</sup>.

BBLR merupakan salah satu faktor utama yang berpengaruh terhadap kematian perinatal dan neonatal. Berdasarkan data dari *The Fifty Sixth Session of Regional Committee WHO for South-East Asia*, pada tahun 2005, kematian bayi terjadi pada usia neonatal dengan penyebab infeksi 33%, asfiksia atau trauma 28%, BBLR 24%, kelainan bawaan 10% dan lain-lain 5%. Risiko kematian BBLR 4 kali lebih besar dibandingkan bayi lahir dengan berat badan lebih dari 2500 gram<sup>[12]</sup>. Bayi berat lahir rendah mempunyai kecenderungan ke arah peningkatan terjadinya infeksi dan mudah terserang komplikasi. Masalah pada BBLR yang sering terjadi adalah gangguan pada sistem pernafasan, susunan saraf pusat, kardiovaskular, hematologi, gastro intestinal, ginjal dan termoregulasi<sup>[6]</sup>. Tingginya angka kejadian BBLR dipengaruhi oleh beberapa faktor antara lain faktor maternal, faktor janin serta faktor lingkungan<sup>[4]</sup>.

Berdasarkan permasalahan tentang berat bayi lahir rendah (BBLR), maka penulis ingin mengidentifikasi dan mengklasifikasi BBL dengan cara memperhatikan data historis ibu hamil di Rumah Sakit Islam Sultan Agung Semarang periode Januari-Desember tahun 2014. Metode klasifikasi yang digunakan dalam penelitian ini adalah *Probabilistic Neural*

*Network* (PNN) dan regresi logistik. Penelitian ini bertujuan untuk mengetahui faktor-faktor yang mempengaruhi berat bayi lahir yang berasal dari faktor ibu (maternal) serta membandingkan antara metode PNN dan regresi logistik dalam pengklasifikasian berat bayi lahir di Rumah Sakit Islam Sultan Agung Semarang tahun 2014.

## 2. TINJAUAN PUSTAKA

### 2.1. Berat Bayi Lahir (BBL)

Berat lahir adalah berat bayi yang ditimbang dalam 1 (satu) jam setelah lahir. Penimbangan ini perlu dilakukan untuk mengetahui apakah berat bayi lahir normal atau rendah. Bayi Berat Lahir Rendah (BBLR) atau *low birth weight infant* didefinisikan sebagai semua berat bayi yang baru lahir dengan berat lahir kurang dari 2.500 gram<sup>[8]</sup>.

### 2.2. Faktor-faktor BBL

Kejadian BBLR dipengaruhi oleh beberapa faktor antara lain faktor maternal, faktor janin serta faktor lingkungan. Faktor maternal antara lain usia ibu, status gizi ibu (tinggi dan berat badan, hemoglobin, tekanan darah dst.), paritas, jarak kehamilan, pendidikan ibu serta penyakit ibu. Faktor janin meliputi hidramnion atau polihidramnion, kehamilan ganda, dan kelainan janin. Sementara faktor lingkungan antara lain fasilitas kesehatan, gaya hidup (perokok, alkohol), serta keadaan sosial ekonomi<sup>[4]</sup>.

### 2.3. Probabilistic Neural Network

*Probabilistic Neural Network* (PNN) merupakan Jaringan Saraf Tiruan (JST) yang dapat digunakan untuk menyelesaikan masalah klasifikasi. Proses yang dilakukan oleh PNN dapat berlangsung lebih cepat bila dibandingkan dengan JST *Back Propagation*. Hal ini disebabkan PNN hanya membutuhkan satu kali iterasi pelatihan bila dibandingkan dengan JST *Back Propagation* yang membutuhkan beberapa kali iterasi pelatihan<sup>[11]</sup>. Walaupun demikian, keakuratan dari klasifikasi PNN sangat ditentukan oleh nilai  $\sigma$  (*smoothing parameter*) dan pola pelatihan yang digunakan. Nilai  $\sigma$  tidak dapat ditentukan secara langsung, tetapi bisa didapatkan melalui metode optimasi atau dari hasil coba-coba.

Misalkan matriks data latih adalah  $\mathbf{M}$ . Algoritma pengujian dengan PNN adalah sebagai berikut<sup>[7]</sup>:

1. Tetapkan: Matriks input  $\mathbf{X}$  berukuran  $q \times p$

Vektor target  $\mathbf{T}$  yang bersesuaian dengan matriks  $\mathbf{M}$ , berukuran  $n$  (terdiri dari  $r$  kelas). Tiap-tiap vektor target ini akan bernilai 1 hanya pada baris yang berhubungan dengan vektor input tertentu, yang lainnya akan bernilai 0. Misal terdapat dua target, maka vektor target bernilai (1, 0) untuk target 1 dan (0, 1) untuk target 2.

2. Inisialisasi: Bobot lapisan input :  $\mathbf{V} = \mathbf{M}$

Bobot bias lapisan input:  $b = \frac{\sqrt{-\ln(0,5)}}{\sigma}$  dengan  $\sigma = (0, 1]$

Bobot lapisan output:  $\mathbf{W} = \mathbf{T}$ .

$$\text{dengan } \mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{q1} & X_{q2} & \cdots & X_{qp} \end{bmatrix}; \mathbf{V} = \begin{bmatrix} V_{11} & V_{12} & \cdots & V_{1p} \\ V_{21} & V_{22} & \cdots & V_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ V_{n1} & V_{n2} & \cdots & V_{np} \end{bmatrix}; \mathbf{W} = \begin{bmatrix} W_{1h} \\ W_{2h} \\ \vdots \\ W_{nh} \end{bmatrix}$$

dimana  $h=1, 2, \dots, r$  (kelas target yang bersesuaian dengan data latih).

3. Kerjakan untuk  $k=1$  sampai  $q$ , pada semua data yang akan diuji:

- a. Cari jarak  $X_k$  terhadap  $V_i$ ,  $i=1, 2, \dots, n$  sebagai berikut:

$$D_{ki} = \sqrt{\sum_{j=1}^p (X_{kj} - V_{ij})^2} \quad (1)$$

b. Cari aktivasi  $a1_{ki}$  sebagai berikut:

$$a1_{ki} = e^{-(b_i D_{ki})^2}; \text{ dengan } i=1, 2, \dots, n. \quad (2)$$

c. Cari  $a2_{kh}$  sebagai berikut:

$$a2_{kh} = \sum_{i=1}^n (a1_{ki} W_{ih}); \text{ dengan } h=1, 2, \dots, r. \quad (3)$$

d. Cari z sedemikian hingga  $a2_{kz} = \max(a2_{kh} | h=1, 2, \dots, r).$

e. Tetapkan z sebagai kelas dari  $X_k$ .

## 2.4. Regresi Logistik

Regresi logistik merupakan metode yang dapat digunakan untuk mencari hubungan antara satu variabel dependen (respon) yang dilambangkan dengan  $y$  yang bersifat *dichotomous* atau *polychotomous* dengan satu atau lebih variabel independen (bebas) yang dilambangkan dengan  $x$ <sup>[1]</sup>.

Hasil observasi variabel acak dependen ( $y$ ) mempunyai dua kategori yaitu 0 dan 1, sehingga mengikuti distribusi Bernoulli dengan fungsi kepadatan peluang<sup>[3]</sup>:

$$P(Y = y) = \pi^y (1 - \pi)^{1-y}; y = 0, 1 \quad (4)$$

dimana jika  $y = 0$  maka  $P(Y = 0) = 1 - \pi$  dan jika  $y = 1$  maka  $P(Y = 1) = \pi$ . Pada regresi logistik ini dapat disusun model yang terdiri dari banyak variabel independen yang dikenal sebagai model multivariabel. Rata-rata bersyarat dari  $y$  jika diberikan nilai  $x$  adalah  $\pi(x) = E(y|x)$ . Model regresi logistik multivariabel dengan  $p$  variabel independen adalah:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)} \quad (5)$$

## 2.5. Penaksir Parameter Regresi Logistik

Untuk menentukan estimasi parameter  $\beta$  digunakan metode *Maximum Likelihood Estimation* (MLE) dengan iterasi Newton Raphson yang merupakan penyelesaian dari turunan pertama dari fungsi likelihood. Penaksiran varian dan kovarian diperoleh dari turunan kedua fungsi ln likelihood. Untuk  $y_i$  berdistribusi binomial, maka fungsi likelihoodnya adalah<sup>[1]</sup>:

$$L(\beta) = \prod_{i=1}^n P(Y = y_i) = \prod_{i=1}^n \pi(x_i)^{y_i} (1 - \pi(x_i))^{n_i - y_i} \quad (6)$$

karena  $\pi(x_i) = \frac{\exp(\sum_{j=0}^p \beta_j x_{ij})}{1 + \exp(\sum_{j=0}^p \beta_j x_{ij})}$  dan  $x_{i0} = 1$ , maka

$$L(\beta) = \left\{ \prod_{i=1}^n \left[ 1 + \exp\left(\sum_{j=0}^p \beta_j x_{ij}\right) \right]^{-n_i} \right\} \exp\left[\sum_{j=0}^p (\sum_{i=1}^n y_i x_{ij}) \beta_j\right]$$

Untuk memudahkan perhitungan, maka fungsi likelihood dimaksimumkan dalam bentuk  $\ln L(\beta)$ .

$$\ln L(\beta) = \ln\left\{ \prod_{i=1}^n \left[ 1 + \exp\left(\sum_{j=0}^p \beta_j x_{ij}\right) \right]^{-n_i} \right\} + \ln\left\{ \exp\left[\sum_{j=0}^p (\sum_{i=1}^n y_i x_{ij}) \beta_j\right] \right\} \quad (7)$$

Menurunkan  $\ln L(\beta)$  terhadap  $\beta_j$  dan hasilnya sama dengan nol

$$\frac{\partial \ln L(\beta)}{\partial \beta_j} = \sum_{i=1}^n y_i x_{ij} - \sum_{i=1}^n n_i x_{ij} \pi(x_i) = 0$$

Penaksiran varian dan kovarian diperoleh dari turunan kedua dari fungsi likelihood<sup>[1]</sup>:

$$\begin{aligned} \frac{\partial^2 \ln L(\beta)}{\partial \beta_j^2} &= - \sum_{i=1}^n n_i x_{ij} x_{ij} \left( \frac{\exp(\sum_{j=0}^p \beta_j x_{ij})}{1 + \exp(\sum_{j=0}^p \beta_j x_{ij})} \right) \left( \frac{1}{1 + \exp(\sum_{j=0}^p \beta_j x_{ij})} \right) \\ \frac{\partial^2 \ln L(\beta)}{\partial \beta_j \partial \beta_{j^*}} &= - \sum_{i=1}^n n_i x_{ij} x_{ij} \pi(x_i) (1 - \pi(x_i)); \quad j, j^* = 1, 2, \dots, p \end{aligned} \quad (8)$$

Nilai parameter  $\beta$  dari turunan pertama fungsi  $L(\beta)$  didapatkan melalui suatu prosedur iterasi menggunakan metode iterasi Newton Raphson hingga didapatkan hasil yang

konvergen yaitu  $|\beta^{(t+1)} - \beta^{(t)}| \leq \varepsilon$ , dimana  $\varepsilon$  bilangan positif yang sangat kecil. Persamaan umum Newton Raphson sebagai berikut:

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - (\mathbf{H}^{(t)}(\boldsymbol{\beta}^{(t)}))^{-1} \mathbf{g}^{(t)}(\boldsymbol{\beta}^{(t)})$$

dengan:  $\boldsymbol{\beta}^{(t)}$  = penaksir parameter yang konvergen pada iterasi ke- $t$ .

$\mathbf{g}^{(t)}(\boldsymbol{\beta}^{(t)})$  = matriks turunan pertama terhadap parameter-parameternya.

$\mathbf{H}^{(t)}(\boldsymbol{\beta}^{(t)})$  = matriks turunan kedua terhadap parameter-parameternya.

## 2.6. Pengujian Parameter

Untuk menguji signifikansi dari parameter dalam model maka dilakukan pengujian secara simultan, parsial serta uji kesesuaian model. Pengujinya adalah sebagai berikut [3]:

### a. Uji Secara Simultan

Untuk menguji signifikansi parameter  $\boldsymbol{\beta}$  dalam model secara simultan menggunakan statistik uji  $G$ .

Hipotesis:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{paling sedikit satu } \beta_j \neq 0; j = 1, 2, \dots, p$$

Taraf signifikansi:  $\alpha$

$$\text{Statistik uji: } G = -2 \ln \left[ \frac{\left( \frac{n_1}{n} \right)^{n_1} \left( \frac{n_0}{n} \right)^{n_0}}{\prod_{i=1}^n (\hat{\pi}(x_i))^{y_i} (1 - (\hat{\pi}(x_i))^{1-y_i})} \right]$$

$$\text{dengan } n_1 = \sum_{i=1}^n y_i; n_0 = \sum_{i=1}^n (1 - y_i); n = n_0 + n_1$$

Kriteria uji: tolak  $H_0$  jika  $G > \chi^2_{(a,p)}$  dengan  $p$  adalah banyaknya variabel independen.

### b. Uji Parsial

Untuk mengetahui signifikansi parameter  $\boldsymbol{\beta}$  terhadap variabel dependennya secara parsial menggunakan statistik uji Wald.

Hipotesis:

$$H_0 : \beta_j = 0; j = 1, 2, \dots, p$$

$$H_1 : \beta_j \neq 0; j = 1, 2, \dots, p$$

Taraf signifikansi:  $\alpha$

$$\text{Statistik uji: } W_j^2 = \frac{\hat{\beta}_j^2}{(SE(\hat{\beta}_j))^2} \text{ dimana } SE(\hat{\beta}_j) = \sqrt{var(\hat{\beta}_j)}$$

Kriteria uji: tolak  $H_0$  jika  $W_j^2 > \chi^2_{(a,1)}$

### c. Uji Kesesuaian Model (*Goodness of Fit*)

Uji kesesuaian model digunakan untuk menilai apakah model sesuai atau tidak menggunakan statistik uji Hosmer dan Lemeshow.

Hipotesis:

$$H_0 : \text{Model sesuai}$$

$$H_1 : \text{Model tidak sesuai}$$

Taraf Signifikansi:  $\alpha$

$$\text{Statistik Uji: } C = \sum_{k=1}^g \left[ \frac{(O_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)} \right] \text{ dengan } O_k = \sum_{i=1}^{n'_k} y_i; \bar{\pi}_k = \sum_{i=1}^{n'_k} \frac{m_i \hat{\pi}_i}{n'_k}$$

dimana:  $O_k$  = observasi pada grup k

$\bar{\pi}_k$  = rata-rata taksiran peluang

$n'_k$  = banyak observasi pada grup k

$g$  = banyaknya grup

$m_i$  = banyaknya observasi yang memiliki nilai  $\hat{\pi}_i$

Kriteria uji: tolak  $H_0$  jika  $C > \chi^2_{(g-2;\alpha)}$

## 2.7. Uji Multikolinieritas

Salah satu cara mengukur multikolinieritas adalah *Variance Inflation Factor (VIF)*. Nilai VIF dapat dihitung berdasarkan nilai koefisien determinasi ( $R^2$ ) dengan cara meregresikan variabel bebas yang satu dengan variabel-variabel bebas yang lain. Semakin tinggi nilai  $R^2$  menunjukkan semakin kuat hubungan linier antar variabel bebas.

$$VIF(X_j) = \frac{1}{(1-R_j^2)} \quad (9)$$

dimana  $R_j^2$  = koefisien determinasi untuk  $X_j$ .

Semakin tinggi VIF semakin berat dampak dari multikolinieritas. Pada umumnya, multikolinieritas dikatakan berat apabila angka VIF  $> 10^{[10]}$ .

## 2.8. Pengukuran Ketepatan Klasifikasi

Salah satu pengukur kinerja klasifikasi adalah tingkat akurasi. Sebuah sistem dalam melakukan klasifikasi diharapkan dapat mengklasifikasi semua set data dengan benar, tetapi tidak dipungkiri bahwa kinerja suatu sistem tidak bisa 100% akurat. Pada umumnya, pengukuran kinerja klasifikasi dilakukan dengan matriks konfusi. Matriks konfusi merupakan tabel pencatat hasil kerja klasifikasi <sup>[9]</sup>.

Akurasi hasil klasifikasi yang dapat dihitung dengan formula sebagai berikut:

- Untuk menghitung akurasi digunakan formula:

$$\text{Akurasi} = \frac{\text{jumlah data yang prediksi benar}}{\text{jumlah prediksi yang dilakukan}} \times 100\%$$

- Untuk menghitung kesalahan prediksi (eror) atau APER (*Apparent Error Rate*) digunakan formula: APER= 1- Akurasi

## 3. METODE PENELITIAN

### 3.1. Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang berasal dari buku status pasien ibu melahirkan yang terdapat pada Rekam Medis Rumah Sakit Islam Sultan Agung Semarang. Jumlah data yang digunakan adalah sebanyak 98 data.

### 3.2. Variabel Penelitian

Penelitian ini menggunakan dua variabel, yaitu berat bayi lahir sebagai variabel respon, sedangkan variabel bebasnya adalah faktor-faktor yang mempengaruhi berat bayi lahir yang berasal dari ibu seperti umur ibu, lama gestasi, tinggi badan, berat badan, tensi darah, hemoglobin serta paritas.

### 3.3. Tahapan Analisis

Program komputer yang digunakan untuk mendukung proses penelitian ini adalah *software Microsoft Excel 2007*, *SPSS 17* dan *Matlab 2009a*. Adapun langkah-langkah pengolahan data yang dilakukan sebagai berikut:

- Merekap data yang akan digunakan dalam penelitian dengan *software Microsoft Excel 2007*
- Melakukan Uji Multikolinieritas antar variabel bebasnya
- Membagi data tersebut menjadi data *training* (data latih) dan data *testing* (data uji)
- Menganalisis klasifikasi menggunakan PNN dengan langkah:
  - Menentukan nilai  $\sigma$  yang akan digunakan dalam PNN
  - Membuat pola PNN menggunakan data latih
  - Menguji pola data uji berdasarkan data latih tersebut

5. Membuat model regresi logistik dan menguji parameter secara simultan dan parsial serta kesesuaian modelnya
6. Menghitung keakuratan klasifikasi baik pada metode PNN maupun regresi logistik
7. Mendapatkan metode terbaik yang memiliki keakuratan terbesar
8. Membuat kesimpulan.

#### 4. HASIL DAN PEMBAHASAN

##### 4.1. Diagnosis Multikolinieritas

Mendiagnosa multikolinieritas dilakukan untuk mengetahui apakah variabel-variabel bebas tidak saling berkorelasi. Untuk mendiagnosa korelasi antar variabel bebas dapat dilihat menggunakan nilai *Variance Inflation Factor* (VIF). Berikut ini nilai VIF masing-masing variabel bebas yang mempengaruhi Berat Bayi Lahir:

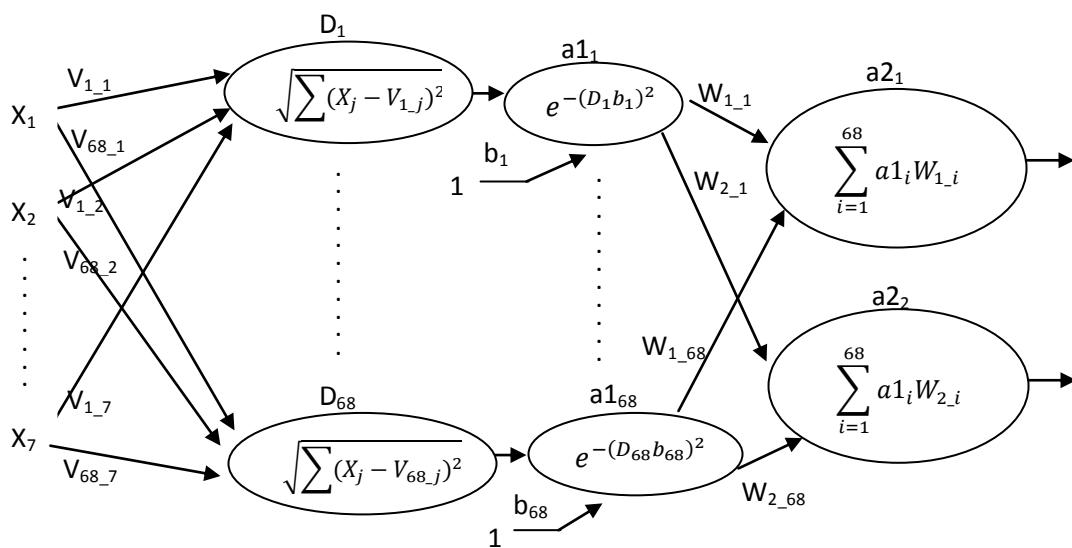
**Tabel 1.** Nilai VIF Variabel Bebas

	X1	X2	X3	X4	X5	X6	X7
R <sup>2</sup>	0,464	0,177	0,219	0,093	0,113	0,195	0,463
VIF	1,867	1,216	1,281	1,102	1,128	1,242	1,863

Tabel 1 menunjukkan bahwa tidak ada variabel bebas yang mengalami multikolinieritas karena semua nilai VIF tiap variabel kurang dari 10, sehingga dapat dikatakan bahwa antar variabel bebas tidak saling berkorelasi atau tidak adanya multikolinieritas diantara variabel-variabel bebas.

##### 4.2. Klasifikasi BBL Menggunakan PNN

Dalam klasifikasi PNN, data dibagi menjadi dua yaitu data latih dan data uji. Dari 98 sampel data yang digunakan, 68 data digunakan sebagai data latih dan 30 data digunakan sebagai data uji. Gambar 1 merupakan arsitektur jaringan PNN yang dibentuk berdasarkan penelitian ini.



**Gambar 1.** Arsitektur Jaringan PNN

Pada Gambar 1 terlihat bahwa Jaringan PNN yang dibangun memiliki jumlah neuron pada lapisan tersembunyi sebanyak 68 neuron dengan 24 neuron berkategori kelas BBLR dan 44 neuron berkategori kelas Normal, serta memiliki 2 neuron pada lapisan outputnya

sesuai dengan jumlah kelasnya yaitu kelas BBLR dan kelas Normal. Sebelum membuat pola pelatihan harus ditentukan nilai penghalus ( $\sigma$ ) terlebih dahulu. Dalam penelitian ini, untuk menentukan nilai penghalus ( $\sigma$ ) dilakukan secara coba-coba dalam selang 0 sampai 1 (0,1] dengan rentang tertentu. Berikut merupakan ketepatan klasifikasi dari PNN dengan beberapa nilai penghalus ( $\sigma$ ):

**Tabel 2.** Ketepatan Klasifikasi PNN

Nilai Penghalus ( $\sigma$ )	Data Latih (%)	Data Uji (%)
0,2	100	46,67
0,4	100	83,33
0,6	100	86,67
0,8	100	86,67
1	100	86,67

Dari Tabel 2 terlihat bahwa ketika menggunakan nilai penghalus sebesar 0,2 sampai 1 ketepatan klasifikasi data latih menghasilkan akurasi sebesar 100%. Hal ini mengindikasikan bahwa pola dari PNN tersebut sangat baik untuk digunakan dalam klasifikasi. Namun, terdapat perbedaan dalam ketepatan klasifikasi data uji pada beberapa nilai penghalus. Sehingga pada penelitian ini, dalam menggunakan pola model PNN akan menghasilkan akurasi yang baik bila menggunakan nilai penghalus dari 0,6 sampai 1.

### 4.3. Pemodelan BBL Menggunakan Regresi Logistik

Data yang digunakan dalam pemodelan regresi logistik ini menggunakan data latih (*training*) yang juga digunakan pada data latih PNN, yaitu sebanyak 68 data, sementara 30 data menjadi data uji (*testing*) yang akan dihitung tingkat ketepatan klasifikasinya berdasarkan model regresi logistik yang diperoleh dari data latih.

Setelah dilakukan olah data menggunakan *Software SPSS 17* didapatkan model awal regresi logistik Berat Bayi Lahir sebagai berikut:

$$\hat{\pi}(x) = \frac{\exp(-44,458 - 0,081X_1 + 1,032X_2 + 0,141X_3 - 0,093X_4 - 0,014X_5 + 1,183X_6 + 0,446X_7)}{1 + \exp(-44,458 - 0,081X_1 + 1,032X_2 + 0,141X_3 - 0,093X_4 - 0,014X_5 + 1,183X_6 + 0,446X_7)}$$

#### 4.3.1. Uji Parameter Model Secara Keseluruhan (Simultan)

Pengujian model regresi logistik secara simultan bertujuan untuk mengetahui signifikansi parameter  $\beta$  terhadap variabel respon secara bersama-sama dengan menggunakan statistik uji  $G$ . Hipotesis yang dilakukan sebagai berikut:

$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$  (secara bersama-sama variabel bebas tidak mempengaruhi model)

$H_1 : \text{paling sedikit satu } \beta_j \neq 0 ; j = 1, 2, \dots, 7$  (secara bersama-sama variabel bebas mempengaruhi model)

**Tabel 3.** Nilai Statistik Uji  $G$  Model Regresi Logistik

Statistik $G$	Df	Sig.
51,560	7	0,000

Berdasarkan Tabel 3 didapatkan nilai  $G$  sebesar 51,560. Dengan menggunakan  $\alpha$  sebesar 5% maka nilai tabel  $Chi-Square \chi^2_{(0,05;7)} = 14,067$ . Dari hasil tersebut dapat dilihat bahwa nilai  $G$  lebih besar dari nilai tabel  $Chi-square$ , berarti menolak  $H_0$  yaitu secara bersama-sama variabel bebas mempengaruhi model.

### 4.3.2. Uji Parameter Model Secara Individu

Pengujian secara individu untuk mengetahui variabel bebas mana saja yang berpengaruh secara signifikan terhadap model. Hipotesis dalam uji parsial yaitu sebagai berikut:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0 ; j = 1, 2, \dots, 7$$

**Tabel 4.** Penaksir Parameter Model Regresi Logistik

Variabel	Koefisien	SE Koefisien	Wald	Sig.	Ket.
X1	-0,081	0,123	0,433	0,511	Tidak
X2	1,032	0,343	9,047	0,003	Signifikan
X3	0,141	0,065	4,664	0,031	Signifikan
X4	-0,093	0,080	1,347	0,246	Tidak
X5	-0,014	0,025	0,349	0,555	Tidak
X6	1,183	0,476	6,179	0,013	Signifikan
X7	0,446	0,585	0,581	0,446	Tidak
Constant	-44,458	18,368	5,858	0,016	

Berdasarkan Tabel 4, apabila digunakan tingkat signifikansi 5% maka parameter yang berpengaruh signifikan terhadap model yang memiliki nilai  $W_j^2$  lebih besar  $\chi^2_{(0,05;1)} = 3,841$  yaitu variabel X<sub>2</sub>, X<sub>3</sub> dan X<sub>6</sub>. Sehingga variabel bebas yang berpengaruh terhadap Berat Bayi Lahir adalah Lama Gestasi, Berat Badan serta Hemoglobin (Hb). Akan tetapi model yang digunakan dalam klasifikasi ini menggunakan semua variabel bebas baik yang signifikan maupun tidak terhadap model.

### 4.3.3. Uji Kecocokan Model (*Goodness of Fit*)

Pada uji kesesuaian model digunakan untuk menilai apakah model sesuai atau tidak. Hipotesis yang digunakan sebagai berikut:

$$H_0 : \text{Model Sesuai}$$

$$H_1 : \text{Model Tidak Sesuai}$$

**Tabel 5.** Nilai Statistik Uji *Goodness of Fit* Model Regresi Logistik

Statistik Uji	Df	Sig.
6,400	8	0,603

Berdasarkan Tabel 5 didapatkan nilai C sebesar 6,400. Nilai C jika dibandingkan dengan nilai  $\chi^2_{(0,05;8)}$  sebesar 15,507 maka nilai statistik C lebih kecil daripada nilai  $\chi^2_{(0,05;8)}$ . Sehingga H<sub>0</sub> diterima dan dapat disimpulkan bahwa model sesuai.

Setelah dilakukan uji signifikansi parameter baik secara simultan maupun parsial serta uji kesesuaian model, maka didapatkan model regresi logistik untuk berat bayi lahir di RSI Sultan Agung Semarang tahun 2014 adalah sebagai berikut:

$$\hat{\pi}(x) = \frac{\exp(-44,458 - 0,081X_1 + 1,032X_2 + 0,141X_3 - 0,093X_4 - 0,014X_5 + 1,183X_6 + 0,446X_7)}{1 + \exp(-44,458 - 0,081X_1 + 1,032X_2 + 0,141X_3 - 0,093X_4 - 0,014X_5 + 1,183X_6 + 0,446X_7)}$$

Berdasarkan model tersebut, variabel yang berpengaruh secara nyata pada model regresi logistik adalah variabel X<sub>2</sub>, X<sub>3</sub> dan X<sub>6</sub>.

#### 4.4.4. Klasifikasi Regresi Logistik

Dengan model regresi yang diperoleh, dapat dihitung nilai ketepatan klasifikasi baik data latih maupun data uji. Untuk mengetahui akurasi klasifikasi dari data uji, dapat dihitung secara manual menggunakan model yang didapat. Apabila didapatkan peluang ( $\pi(x)$ ) < 0,5 maka data tersebut dikategorikan BBLR, sedangkan apabila memiliki peluang  $\geq 0,5$  maka akan dikategorikan BBL Normal.

**Tabel 6.** Klasifikasi Data Latih Menggunakan Regresi Logistik

Observasi	Prediksi		Akurasi (%)
	BBLR	Normal	
<b>BBLR</b>	20	4	83,3
<b>Normal</b>	4	40	90,9
<b>Total</b>			88,2

Berdasar Tabel 6, terlihat bahwa akurasi ketepatan klasifikasi data latih sebesar 88,2%. Dimana terjadi kesalahan klasifikasi pada BBL kategori BBLR yang diidentifikasi berkategori Normal sebanyak 4 data, sedangkan 4 data kategori Normal yang diidentifikasi berkategori BBLR.

**Tabel 7.** Klasifikasi Data Uji Menggunakan Regresi Logistik

Observasi	Prediksi		Akurasi (%)
	BBLR	Normal	
<b>BBLR</b>	7	3	70
<b>Normal</b>	1	19	95
<b>Total</b>			86,67

Berdasar Tabel 7, terlihat akurasi ketepatan klasifikasi data uji dengan menggunakan regresi logistik sebesar 86,67%. Dimana untuk kategori BBLR akurasinya sebesar 70% dengan 3 data salah identifikasi, serta untuk kategori Normal akurasinya sebesar 95% dengan satu data yang salah identifikasi masuk kategori BBLR.

#### 4.4. Perbandingan Metode PNN dengan Metode Regresi Logistik

Perbandingan metode PNN dengan metode regresi logistik dilakukan setelah didapatkan pemodelan Berat Bayi Lahir di RSI Sultan Agung Semarang tahun 2014 dengan *Probabilistik Neural Network* maupun regresi logistik. Perbandingan ini bertujuan untuk mengetahui metode mana yang lebih tepat pengklasifikasianya. Hasil yang diperoleh dapat dilihat pada Tabel 8.

**Tabel 8.** Perbandingan Ketepatan Klasifikasi

Model	Ketepatan Klasifikasi (%)	
	Data Latih	Data Uji
PNN	100	86,67
Regresi Logistik	88,2	86,67

Tabel 8 menunjukkan bahwa klasifikasi Berat Bayi Lahir menggunakan metode *Probabilistik Neural Network* menghasilkan nilai akurasi klasifikasi yang lebih besar dibandingkan dengan metode regresi logistik pada data latih, sedangkan pada data uji

memiliki nilai akurasi yang sama besar. Hal ini menunjukkan bahwa metode *Probabilistik Neural Network* lebih baik apabila digunakan pada klasifikasi Berat Bayi Lahir di RSI Sultan Agung Semarang tahun 2014.

## 5. KESIMPULAN

Berdasarkan hasil dan pembahasan, dapat diperoleh kesimpulan sebagai berikut:

1. Metode PNN terbaik adalah dengan menggunakan  $\sigma$  sebesar 0,6; 0,8 dan 1 karena memiliki akurasi 100% pada data latih dan 86,67% pada data ujinya.
2. Model regresi logistik Berat Bayi Lahir di RSI Sultan Agung Semarang tahun 2014 yaitu:

$$\hat{\pi}(x) = \frac{\exp(-44,458 - 0,081X_1 + 1,032X_2 + 0,141X_3 - 0,093X_4 - 0,014X_5 + 1,183X_6 + 0,446X_7)}{1 + \exp(-44,458 - 0,081X_1 + 1,032X_2 + 0,141X_3 - 0,093X_4 - 0,014X_5 + 1,183X_6 + 0,446X_7)}$$

Variabel bebas yang mempengaruhi secara signifikan terhadap Berat Bayi Lahir di RSI Sultan Agung Semarang tahun 2014 yaitu variabel Lama Gestasi ( $X_2$ ), Berat badan ( $X_3$ ) dan Hemoglobin ( $X_6$ ).

3. Pemodelan Berat Bayi Lahir di RSI Sultan Agung Semarang tahun 2014 menggunakan metode *Probabilistik Neural Network* lebih baik daripada menggunakan metode regresi logistik, dimana ketepatan klasifikasi metode PNN sebesar 100% pada data latih dan 86,67% pada data ujinya.

## 6. DAFTAR PUSTAKA

- [1] Agresti, A. 2002. *Categorical Data Analysis, Second Edition*. New York: John Wiley & Sons.
- [2] Azikin, G. 2011. *Gambaran Kejadian Bayi Berat Lahir Rendah (BBLR) di Rumah Sakit Umum*. [http://kebidanan-kti.blogspot.com/2011/11/gambaran-kejadian-bayi-berat-lahir\\_19.html](http://kebidanan-kti.blogspot.com/2011/11/gambaran-kejadian-bayi-berat-lahir_19.html), (diakses Selasa, 10 Maret 2015 pukul 16.00 WIB).
- [3] Hosmer, D. W. and Lemeshow, S. 2000. *Applied Logistic Regression, Second Edition*. New York: John Wiley & Sons.
- [4] Jaya, N. 2009. *Analisis Faktor Resiko Kejadian Bayi Berat Lahir Rendah di Rumah Sakit Ibu dan Anak Siti Fatimah Kota Makassar*. Media Gizi Pangan Vol. VII Edisi 1: 49-54.
- [5] Kemenkes RI. 2013. *Profil Kesehatan Indonesia 2012*. Jakarta: Kementerian Kesehatan Republik Indonesia.
- [6] Kemenkes RI. 2014. *Profil Kesehatan Indonesia 2013*. Jakarta: Kementerian Kesehatan Republik Indonesia.
- [7] Kusumadewi, S. 2004. *Klasifikasi Pola Menggunakan Jaringan Probabilistik*. Seminar Nasional Aplikasi Teknologi Informasi Hal: J59-J64.
- [8] Novita, R.V.T. 2011. *Keperawatan Maternitas*. Bogor: Ghalia Indonesia.
- [9] Prasetyo, E. 2012. *Data Mining :Konsep dan aplikasi Menggunakan Matlab*. Yogyakarta: Andi Publisher.
- [10] Sarwoko. 2005. *Dasar-dasar Ekonometrika*. Yogyakarta: Andi.
- [11] Specht, D.F. 1990. *Probabilistic Neural Networks*. Neural Networks Vol. 3: 109-118.
- [12] Yulifah, R. dan Yuswanto, T.J.A. 2009. *Asuhan Kebidanan Komunitas*. Jakarta: Salemba Medika.