

**KETEPATAN KLASIFIKASI KEIKUTSERTAAN KELUARGA BERENCANA  
(KB) MENGGUNAKAN ANALISIS REGRESI LOGISTIK BINER DAN  
FUZZY K-NEAREST NEIGHBOR IN EVERY CLASS  
DI KABUPATEN KLATEN**

**Dhinda Amalia Timur<sup>1</sup>, Yuciana Wilandari<sup>2</sup>, Diah Safitri<sup>3</sup>**

<sup>1</sup>Mahasiswa Jurusan Statistika FSM UNDIP

<sup>2,3</sup>Staf Pengajar Jurusan Statistika FSM UNDIP

**ABSTRACT**

Fertility is one of the factors that affect population growth. High population growth resulted in the emergence of a variety of problems for a country including Indonesia. This requires a treatment that population growth can be controlled, one attempts to handle by using a Keluarga Berencana program. Therefore conducted a study to determine the factors that affect that participation of Keluarga Berencana (KB) by using Binary Logistic Regression analysis in which the participation of KB divided into two, namely join KB and KB did not participate. Based on the results obtained Binary logistic regression analysis predictor variables that significantly affect participation KB is the number of children, father's education, and mother's education. The resulting classification accuracy with training data comparison testing was 90:10 at 84.375%. Furthermore, the data were analyzed by using Fuzzy K-Nearest Neighbor in every Class (FK-NNC) to determine the accuracy of the classification results comparison with FK-NNC Binary Logistic Regression. From the analysis of the classification accuracy using the FK-NNC with a 90:10 ratio of training data and testing the value of  $K = 7$  values obtained terbesar ie 87.5%. The comparison of classification accuracy of this value indicates if the FK-NNC is better classify participation in Keluarga Berencana in Klaten district 2012.

**Keywords:** Keluarga Berencana, Binary Logistic Regression, Fuzzy K-Nearest Neighbor in every Class (FK-NNC)

## **1. PENDAHULUAN**

### **1.1. Latar Belakang**

Menurut publikasi BPS pada bulan Agustus 2010, jumlah penduduk Indonesia berdasarkan hasil sensus adalah sebanyak 237.556.363 orang. Diketahui juga dari hasil sensus penduduk 2010 bahwa laju pertumbuhan penduduk sebesar 1,49 persen per tahun. Penduduk Indonesia terus bertambah dari waktu ke waktu. Dari pemerintah Hindia Belanda tahun 1930 hingga tahun 2010 (BPS, 2010).

Fertilisasi merupakan salah satu faktor yang mempengaruhi pertumbuhan penduduk. Tingkat fertilitas yang tinggi merupakan masalah yang serius bagi suatu negara tak terkecuali Indonesia. Pertumbuhan penduduk yang tinggi mengakibatkan timbulnya berbagai masalah. Salah satu cara yang dilakukan oleh pemerintah dalam mengendalikan fertilitas adalah dengan melalui program Keluarga Berencana (KB), program KB dapat dilakukan melalui pemakaian alat kontrasepsi oleh pasangan usia subur (PUS) (BKKBN, 2009)

Beberapa penelitian yang telah dilakukan mengenai KB dalam bidang kesehatan adalah sebagai berikut, Haloho, et al, (2013) Berdasarkan penelitian tersebut diduga terdapat banyak faktor yang mempengaruhi pemakaian alat kontrasepsi. Setelah dilakukan analisis variabel yang berpengaruh signifikan terhadap pemakaian alat

kontrasepsi wanita adalah umur ibu, umur anak terakhir, dan pernah tidaknya mendapatkan penyuluhan terhadap keluarga berencana dari pihak berwenang. Di dalam [www.bkkbn.go.id](http://www.bkkbn.go.id) terdapat penelitian yang sama dibidang kesehatan dilakukan oleh BKKBN, dari hasil penelitian diperoleh kesimpulan bahwa faktor yang berpengaruh terhadap anggota kelompok UPPKS adalah status kesejahteraan, jumlah anak dan kunjungan petugas.

Menurut Agresti (2002), salah satu cara yang digunakan untuk menganalisis hubungan antara suatu variabel dependen dan beberapa variabel independen, dengan variabel dependennya berupa data kualitatif dikotomi yaitu bernilai 1 atau bernilai 0 adalah analisis regresi logistik biner. Selain itu dengan analisis regresi logistik biner dapat diketahui pula nilai ketepatan klasifikasi Selain Regresi Logistik Biner untuk mengetahui ketepatan klasifikasi atau akurasi dapat juga digunakan metode data mining. Salah satu metode tersebut adalah *Fuzzy K-Nearest Neighbor* (FK-NNC). Penelitian sebelumnya mengenai *Fuzzy K-Nearest Neighbor in every Class*, yaitu penelitian Eko Prasetyo (2012b) di bidang kesehatan, dalam penelitiannya Eko Prasetyo membandingkan nilai ketepatan klasifikasi atau akurasi antara metode *K-Nearest Neighbor* (K-NN), *Fuzzy K-Nearest Neighbor* (FK-NN), *Fuzzy K-Nearest Neighbor in every Class* (FK-NNC). Kesimpulan yang didapat adalah akurasi yang diberikan oleh *Fuzzy K-Nearest Neighbor in every Class* (FK-NNC) lebih tinggi daripada K-NN dan FK-NN. Berdasarkan latar belakang tersebut Sehingga pada penelitian ini penulis memilih judul “Ketepatan Klasifikasi Keikutsertaan Keluarga Berencana Menggunakan Analisis Regresi Logistik Biner Dan *Fuzzy K-Nearest Neighbor in every Class* di Kabupaten Klaten”

## 1.2. Tujuan Penelitian

Berdasarkan latar belakang, tujuan penelitian tugas akhir ini adalah:

1. Menentukan faktor-faktor yang mempengaruhi keikutsertaan program Keluarga Berencana berdasarkan analisis regresi logistik
2. Menentukan ketepatan klasifikasi peserta keluarga berencana berdasarkan analisis regresi logistik biner
3. Menentukan ketepatan klasifikasi peserta keluarga berencana berdasarkan analisis *Fuzzy K-Nearest Neighbor in every Class* (FK-NNC)
4. Membandingkan tingkat ketepatan klasifikasi peserta keluarga berencana berdasarkan analisis regresi logistik biner dan *Fuzzy K-Nearest Neighbor in every Class* (FK-NNC)

## 2. TINJAUAN PUSTAKA

### 2.1. Keluarga Berencana

Definisi keluarga berencana menurut Entjang (1986) Keluarga Berencana (KB) adalah daya upaya manusia untuk mengatur, secara sengaja kehamilan dalam keluarga secara tidak melawan hukum dan moral pancasila, demi untuk kesejahteraan keluarga. Hartanto (1994) mengatakan tujuan KB adalah mewujudkan Norma keluarga Kecil yang Bahagia dan Sejahtera (NKKBS). Menurut BKKBN (2011) Peserta KB adalah pasangan usia subur yang suami/isterinya sedang memakai atau menggunakan salah satu alat/cara kontrasepsi modern pada tahun pelaksanaan pendataan keluarga/pemutakhiran data keluarga, sedangkan Pasangan Usia Subur (PUS) yang bukan peserta KB terdiri dari hamil, ingin anak segera, ingin anak tunda, dan tidak ingin anak lagi.

## 2.2. Model Regresi Logistik Biner

Menurut Agresti (2002) Regresi logistik merupakan metode yang menghubungkan antara variabel dependen yang bersifat kategorik dengan variabel independen. Model regresi logistik dengan variabel dependen yang terdiri dari dua kategori disebut model regresi logistik biner (dikotomus).

Menurut Agresti (2002) variabel dependen ( $y$ ) dikatakan sebagai variabel indikator dan memenuhi distribusi Bernoulli. Distribusi Bernoulli yaitu distribusi dari variabel acak yang hanya mempunyai dua kategorik. Fungsi probabilitas untuk setiap observasi sebagai berikut:

$$f(y) = \pi(\mathbf{x})^y(1 - \pi)^{1-y}, \quad y = 0,1$$

dimana jika  $y = 0$  maka  $f(y) = 1 - \pi(\mathbf{x})$  dan jika  $y = 1$  maka  $f(y) = \pi(\mathbf{x})$ . Fungsi regresi logistiknya dapat dituliskan:

$$f(x) = \frac{1}{1+e^{-z}} \text{ ekuivalen } f(x) = \frac{e^z}{1+e^z}$$

dengan:  $z = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_px_p$

Menurut Hosmer dan Lemeshow (2000), bentuk model regresi logistik dengan variabel independen  $x$  adalah sebagai berikut:

$$\pi(\mathbf{x}) = \frac{e^{\beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_px_p}}{1 + e^{\beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_px_p}}$$

Sedangkan logit dari  $\pi(x)$  adalah

$$g(\mathbf{x}) = \ln\left(\frac{\pi(\mathbf{x})}{1-\pi(\mathbf{x})}\right) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_px_p$$

Jika  $(\mathbf{x}) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_px_p$ ,  $g(\mathbf{x})$  merupakan fungsi hubungan dari model regresi logistik yang disebut fungsi hubungan logit.

## 2.3. Estimasi Parameter

Menurut Agresti (2002) untuk menentukan estimasi parameter digunakan metode Maksimum Likelihood yang membutuhkan turunan pertama dan turunan kedua dari fungsi likelihood.

Karena setiap pengamatan diasumsikan independen maka Hosmer dan Lemeshow (2000) menyatakan fungsi likelihoodnya merupakan perkalian antara masing-masing fungsi densitas yaitu:

$$l(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i}$$

Pada prinsipnya, fungsi maksimum likelihood menggunakan estimasi nilai  $\beta$  untuk memaksimalkan fungsi likelihood, dan logaritma likelihoodnya menurut Hosmer dan Lemeshow (2000) dinyatakan sebagai:

$$L(\beta) = \ln[l(\beta)] = \sum_{i=1}^n [y_i \ln\{\pi(x_i)\} + (1 - y_i) \ln\{1 - \pi(x_i)\}]$$

Hosmer dan Lemeshow (2000) menyatakan untuk memperoleh nilai  $\beta$  maka dengan memaksimumkan  $L(\beta)$  dan mendiferensialkan  $L(\beta)$  terhadap  $\beta_0$  dan  $\beta_j$  dengan  $j = 1, 2, \dots, p$  dan menyamakannya dengan 0, persamaan ini dapat ditulis dalam bentuk sebagai berikut:

$$\sum_{i=1}^n (y_i - \pi(x_i)) = 0 \text{ dan } \sum_{i=1}^n (x_i [y_i - \pi(x_i)]) = 0$$

Hasil turunan pertama dalam bentuk matriks adalah  $X'(Y - \pi(x))$ , dengan:

$$X' = \begin{bmatrix} 1 & 1 & \cdot & \cdot & \cdot & 1 \\ x_{11} & x_{21} & \cdot & \cdot & \cdot & x_{n1} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ x_{1p} & x_{2p} & \cdot & \cdot & \cdot & x_{np} \end{bmatrix} \quad (Y - \pi(x)) = \begin{bmatrix} y_1 - \pi(x_1) \\ y_2 - \pi(x_2) \\ \cdot \\ \cdot \\ \cdot \\ y_n - \pi(x_n) \end{bmatrix}$$

Di dalam Hosmer dan Lemeshow (2000) metode untuk mencari estimasi varian kovarian dari koefisien estimasi mengikuti teori yang dikembangkan dengan teori estimasi maksimum likelihood. Teori ini diketahui bahwa estimator diperoleh dari matriks turunan parsial kedua dari fungsi log likelihood. Menurut Hosmer dan Lemeshow (2000) turunan parsial memiliki bentuk umum berikut:

$$\frac{\partial^2 L(\beta)}{(\partial \beta_j)^2} = -\sum_{i=1}^n (x_{ji})^2 \pi_i [1 - \pi] \text{ dan } \frac{\partial^2 L(\beta)}{\partial \beta_u \partial \beta_j} = -\sum_{i=1}^n x_{ui} x_{ji} \pi_i [1 - \pi]$$

Hosmer dan Lemeshow (2000) mengatakan jika dinyatakan dalam bentuk matriks turunan parsial kedua dari log likelihood adalah  $X'VX$ , dengan:

$$V = \begin{bmatrix} \pi(x_1)[1 - \pi(x_1)] & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \pi(x_2)[1 - \pi(x_2)] & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & \pi(x_n)[1 - \pi(x_n)] \end{bmatrix}$$

Menurut Agresti (2002), metode Newton Raphson merupakan metode iteratif untuk memecahkan persamaan yang secara analitik tidak dapat langsung diketahui hasilnya:

1. Dipilih taksiran awal untuk  $\beta$ , misal  $\hat{\beta}^{(1)} = 0$
2. Dihitung  $X'(Y - \pi(x))$  dan  $X'VX$ , selanjutnya dihitung invers dari  $X'VX$
3. Pada setiap  $r + 1$  dihitung taksiran baru yaitu

$$\hat{\beta}^{(r+1)} = \hat{\beta}^{(r)} + \{X'VX\}^{-1} \{X'(Y - \pi(x))\}$$

Iterasi berakhir jika diperoleh  $\hat{\beta}^{(r+1)} \cong \hat{\beta}^{(r)}$

## 2.4. Uji Signifikansi

### 2.5.1 Uji Rasio Likelihood (Uji Keseluruhan)

Hosmer dan Lemeshow (2000) mengatakan uji Rasio Likelihood bertujuan untuk mengetahui apakah variabel independen yang terdapat dalam model berpengaruh

nyata atau tidak secara keseluruhan. Menurut Hosmer dan Lemeshow (2000), langkah-langkah dalam melakukan uji rasio likelihood yaitu:

- a. Hipotesa:  
 $H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$   
 $H_1 : \text{paling sedikit salah satu dari } \beta_j \neq 0 \text{ dengan } j = 1, 2, \dots, p$
- b. Statistik Uji  

$$G = -2 \ln \left( \frac{\text{likelihood tanpa variabel prediktor}}{\text{likelihood dengan variabel prediktor}} \right)$$
- c. Kriteria Uji: Tolak  $H_0$  jika nilai  $G > \chi^2_{(\alpha, p)}$

### 2.5.2 Uji Wald

Menurut Hosmer dan Lemeshow (2000) uji Wald atau uji individu diperoleh dengan cara mengkuadratkan rasio estimasi parameter dengan estimasi standar errornya, uji wald dilakukan untuk mengetahui signifikansi parameter terhadap variabel dependen, langkah-langkah uji Wald sebagai berikut:

- a. Hipotesa:  
 $H_0 : \beta_j = 0 \text{ dengan } j=1, 2, \dots, p$   
 $H_1 : \beta_j \neq 0 \text{ dengan } j=1, 2, \dots, p$
- b. Statistik Uji  

$$W_j = \left\{ \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \right\}^2$$
- c. Kriteria Uji: Tolak  $H_0$  jika nilai  $W_j > \chi^2_{(\alpha, 1)}$

### 2.5.3 Uji Kesesuaian Model

Hosmer dan Lemeshow (2000) mengatakan uji kesesuaian model digunakan untuk menilai apakah model sesuai atau tidak, statistik uji yang digunakan adalah uji Hosmer dan Lemeshow dengan langkah-langkah sebagai berikut:

- a. Hipotesis  
 $H_0 : \text{Model sesuai (tidak ada perbedaan antara hasil observasi dengan hasil prediksi)}$   
 $H_1 : \text{Model tidak sesuai (ada perbedaan antara hasil observasi dengan hasil prediksi)}$
- b. Statistik Uji  

$$\hat{C} = \sum_{f=1}^g \left[ \frac{(O_{1f} - n'_f \bar{\pi}_f)^2}{(n'_f \bar{\pi}_f (1 - \bar{\pi}_f))} \right]$$
  
  
 Dimana:  
 $O_f = \text{observasi pada grup ke } f$                        $n'_f = \text{banyak observasi pada grup ke- } f$   
 $\bar{\pi}_f = \text{rata-rata taksiran peluang}$                        $g = \text{banyak grup}$
- c. Kriteria Uji:  $H_0$  ditolak jika:  $\hat{C} > \chi^2_{(g-2, \alpha)}$

### 2.5. Fuzzy K-Nearest Neighbor in Every Class

Menurut Prasetyo (2012a) metode *Fuzzy K-Nearest Neighbor in Every Class* (FK-NNC) menggunakan sejumlah K tetangga terdekat pada setiap kelas dari sebuah

data uji. Kerangka kerja FK-NNC menggunakan FK-NN sebagai basis kerangka kerja, di mana sebuah data uji mempunyai nilai keanggotaan pada setiap kelas dalam interval [0,1]. Berikut persamaan nilai keanggotaan

$$\sum_{k=1}^C u_{ik} = 1, 0 \leq u_{ik} \leq 1$$

$u_{ik}$  adalah nilai keanggotaan data uji  $x_i$  ke kelas -  $k$

Setiap data uji  $x_i$  harus dicarikan K tetangga terdekat pada setiap kelas menggunakan formula jarak seperti berikut:

$$d(x_i, x_l) = \left( \sum_{j=1}^p |x_{ij} - x_{lj}|^s \right)^{\frac{1}{s}}$$

dimana  $p$  adalah dimensi (jumlah fitur) data. Untuk  $s$  adalah penentu jarak yang digunakan, jika  $s=2$  maka jarak yang digunakan adalah Euclidean.

Menurut Prasetyo (2012a), jarak atau ukuran ketidakmiripan suatu data kategorik ordinal digunakan rumus sebagai berikut:

$$d = |x_{ij} - x_{lj}| / (q - 1)$$

nilainya dipetakan ke tipe integer 0 sampai  $q - 1$ , dimana  $q$  adalah jumlah nilai kategorik

Sedangkan ukuran ketidakmiripan suatu data rasio adalah:

$$d = |x_{ij} - x_{lj}|$$

Selanjutnya menurut Prasetyo (2012a), jarak data uji  $x_i$  ke semua K tetangga dari setiap kelas ke- $k$  dijumlahkan. Formula yang digunakan adalah

$$S_{ik} = \sum_{r=1}^K d(x_i, x_r)^{\frac{-2}{m-1}}$$

Nilai  $d$  sebagai akumulasi jarak data uji  $x_i$  ke K tetangga dalam kelas ke- $k$  dilakukan sebanyak C kelas. Nilai  $m$  di sini merupakan pangkat bobot yang menunjukkan banyak kelas (*weight exponent*). Selanjutnya, akumulasi jarak data uji  $x_i$  ke setiap kelas digabungkan, disimbolkan D. Formula yang digunakan adalah

$$D_i = \sum_{k=1}^C (S_{ik})$$

Untuk mendapatkan nilai  $u_{ik}$  (nilai keanggotaan data uji  $x_i$  pada setiap kelas ke- $k$  (ada C kelas)), dengan menggunakan rumus:

$$u_{ik} = \frac{S_{ik}}{D_i}$$

Untuk menentukan kelas hasil prediksi data uji  $x_i$ , dipilih kelas dengan nilai keanggotaan terbesar dari data  $x_i$ . Formula yang digunakan adalah

$$y' = \arg \max_{k=1}^c (u_{ik})$$

dengan:  $y'$  = kelas prediksi       $c$  = banyak kelas  
 $k = 1, 2$

## 2.6. Ketepatan Klasifikasi

Menurut Johnson dan Wichern (1992) prosedur klasifikasi yang dipakai pada regresi logistik adalah *apparent error rate* (APER), nilai APER menyatakan proporsi sampel yang salah diklasifikasikan oleh fungsi klasifikasi, dengan rumus sebagai berikut:

**Tabel 1. Matriks Konfusi**

Hasil Observasi	Taksiran	
	y1	y2
y1	n <sub>11</sub>	n <sub>12</sub>
y2	n <sub>21</sub>	n <sub>22</sub>

Maka nilai APER dapat dihitung dengan rumus:

$$APER = \frac{n_{12} + n_{21}}{n_{11} + n_{12} + n_{21} + n_{22}}$$

### 3. METODOLOGI PENELITIAN

Sumber data yang digunakan dalam penelitian ini adalah data sekunder yaitu formulir pemutakhiran data keluarga (MDK) kabupaten Klaten tahun 2012.

Variabel penelitian yang dianalisis terdiri dari variabel dependen dan variabel independen. Variabel dependen (Y) yang dianalisis adalah keikutsertaan program KB yang dikategorikan menjadi 2 yaitu:

Y = 0, untuk tidak ikut KB      Y = 1, untuk ikut KB

Sedangkan, variabel independen yang diasumsikan memiliki pengaruh terhadap keikutsertaan program KB adalah :

- a. Umur ibu
- b. Umur anak terakhir
- c. Jumlah anak hidup
- d. Usia menikah
- e. Pendidikan ayah, dikategorikan menjadi tiga:
  - 1 : Pendidikan rendah yaitu setinggi-tingginya pendidikan tingkat SMP
  - 2 : Pendidikan sedang yaitu pendidikan SMA
  - 3 : Pendidikan tinggi yaitu Perguruan Tinggi keatas
- f. Pendidikan ibu, dikategorikan menjadi tiga:
  - 1 : Pendidikan rendah yaitu setinggi-tingginya pendidikan tingkat SMP
  - 2 : Pendidikan sedang yaitu pendidikan SMA
  - 3 : Pendidikan tinggi yaitu Perguruan Tinggi keatas

Tingkat Kesejahteraan

- 1 : Pra Sejahtera
- 2 : KS I
- 3. KS II
- 4. KS III
- 5. KS III+

### 4. ANALISIS DAN PEMBAHASAN

#### 4.1. Analisis Regresi Logistik Biner

Dalam penelitian ini perbandingan data training dan testing yang digunakan adalah 90:10 dan diolah dengan menggunakan SPSS 19. Dalam penelitian ini uji signifikansi parameter dilakukan dua kali hingga ditemukan semua variabel X berpengaruh secara signifikan terhadap variabel Y.

##### 4.1.1. Uji Rasio Likelihood

**Tabel 2. Uji Rasio Likelihood**

Uji ke-	G	Nilai Tabel Chi-Square	Keputusan
1	79, 167	$\chi^2_{(0,05, 12)} = 21,03$	H <sub>0</sub> ditolak
2	77, 577	$\chi^2_{(0,05, 5)} = 11, 07$	H <sub>0</sub> ditolak

Berdasarkan Tabel 2 diketahui bahwa uji rasio likelihood menghasilkan keputusan  $H_0$  ditolak karena nilai  $G > \chi^2_{(0,05, p)}$ . Sehingga dapat disimpulkan bahwa variabel independen yang berpengaruh nyata secara serentak.

#### 4.1.2. Uji Wald

**Tabel 3.** Uji Wald Pertama

Variabel	Wald	Sig.	Keputusan
x1	0,257	0,612	$H_0$ diterima
x2	0,097	0,755	$H_0$ diterima
x3	43,526	0,000	$H_0$ ditolak
x4	0,221	0,639	$H_0$ diterima
x5(1)	4,336	0,037	$H_0$ ditolak
x5(2)	3,483	0,062	$H_0$ diterima
x6 x6(1)	8,602	0,003	$H_0$ ditolak
x6(2)	5,974	0,015	$H_0$ ditolak
x7 x7(1)	0,148	0,701	$H_0$ diterima
x7(2)	0,076	0,783	$H_0$ diterima
x7(3)	0,033	0,857	$H_0$ diterima
x7(4)	0	0,993	$H_0$ diterima

**Tabel 4.** Uji Wald Kedua

Variabel	Wald	Sig	Keputusan
x3	44,697	0,000	$H_0$ ditolak
x5 x5(1)	4,788	0,029	$H_0$ ditolak
x5(2)	3,54	0,06	$H_0$ diterima
x6 x6(1)	9,289	0,002	$H_0$ ditolak
x6(2)	6,73	0,009	$H_0$ ditolak

Keterangan:

$X_1$	= Umur ibu	$X_5$	= Pendidikan ayah
$X_2$	= Umur anak terakhir	$X_6$	= Pendidikan ibu
$X_3$	= Jumlah anak	$X_7$	= Tingkat kesejahteraan
$X_4$	= Umur menikah		

Berdasarkan Tabel 3, didapatkan hasil bahwa variabel umur ibu, umur anak terakhir, umur menikah dan tingkat kesejahteraan dalam rumah tangga, hal ini ditunjukkan oleh nilai  $W_j < \chi^2_{(0,05,1)} = 3,84$  atau nilai sig.  $> \alpha = 0,05$ . Sedangkan variabel jumlah anak, pendidikan ayah dan pendidikan ibu berpengaruh signifikan terhadap variabel dependen keikutsertaan KB, hal ini ditunjukkan oleh nilai  $W_j > \chi^2_{(0,05,1)} = 3,84$  atau nilai sig.  $< \alpha = 0,05$ . Sedangkan Tabel 4 menunjukkan bahwa variabel jumlah anak, pendidikan ayah dan pendidikan ibu berpengaruh terhadap variabel dependen keikutsertaan KB.

#### 4.1.3. Uji Kesesuaian Model

Berdasarkan uji kesesuaian model diperoleh hasil bahwa  $H_0$  diterima, hal ini ditunjukkan dengan nilai  $\hat{C} (7, 344) < \chi^2_{(0,05, 1)} (15,51)$  atau nilai sig.  $(0,500) > \alpha = 0,05$ . Jadi, model regresi logistik biner yang terbentuk sudah sesuai, atau tidak ada perbedaan antara observasi dengan hasil prediksi.

#### 4.1.4. Model Akhir

Setelah dilakukan uji signifikansi terhadap model, baik secara keseluruhan maupun individual. Diperoleh model akhir sebagai berikut:

$$\pi(x) = \frac{\exp(-1,950 + 1,559 x_3 - 1,600 x_5(1) - 1,210 x_5(2) + 2,086 x_6(1) + 1,611 x_6(2))}{1 + \exp(-1,950 + 1,559 x_3 - 1,600 x_5(1) - 1,210 x_5(2) + 2,086 x_6(1) + 1,611 x_6(2))}$$



#### 4. 2. Ketepatan Klasifikasi Menggunakan Regresi Logistik Biner

Dalam menghitung nilai ketepatan klasifikasi model digunakan data testing sebanyak 10% sesuai proporsi data training:testing 90:10

**Tabel 5.** Ketepatan Klasifikasi Model

Observasi	Tidak ikut KB	ikut KB
Tidak ikut KB	9	3
ikut KB	2	18

Dari Tabel 5 dapat dihitung nilai ketepatan klasifikasinya adalah 84,375%, sehingga nilai misklasifikasi (APER) adalah 15,625%. Kemudian disimpulkan bahwa model regresi logistik biner cukup baik dalam mengklasifikasikan keikutsertaan KB.

#### 4. 3. Ketepatan Klasifikasi Menggunakan FK-NNC

Nilai ketepatan klasifikasi dengan menggunakan metode FK-NNC dengan proporsi data training:testing 90:10 dilakukan sebanyak beberapa kali dengan nilai K yang berbeda, nilai K yang digunakan adalah 3,5,7,9,dan 11. Berdasarkan proses perhitungan diperoleh nilai ketepatan klasifikasi sebagai berikut:

**Tabel 6** Hasil Ketepatan Klasifikasi FK-NNC

K	Ketepatan Klasifikasi FK-NNC	APER
3	53,13%	46,88%
5	59,38%	40,62%
7	87,5%	12,5%
9	87,5%	12,5%
11	87,5%	12,5%

diperoleh nilai akurasi tertinggi yaitu dengan nilai K= 7, 9, dan 11 sebesar 87,5% dan APER = 12,5% , dengan nilai laju error sebesar 12,5%.

#### 4. 4. Perbandingan Ketepatan Klasifikasi

Setelah dilakukan perbandingan, diketahui bahwa nilai ketepatan klasifikasi dengan menggunakan Regresi Logistik Biner diperoleh nilai 84,375% dan FK-NNC diperoleh nilai 87,5% Dilihat dari nilai ketepatan klasifikasi pada setiap perbandingan data secara umum menunjukkan bahwa nilai ketepatan klasifikasi dengan menggunakan FK-NNC lebih besar daripada menggunakan metode Regresi Logistik Biner.

**Tabel 7.** Perbandingan Ketepatan Klasifikasi

Training : Testing	Regresi Logisitik Biner	FK-NNC
90 : 10	84,38%	87,50%
80 : 20	83,08%	84,62%
70 : 30	71,43%	74,49%
60 : 40	75,19%	79,07%
50 : 50	73,29%	71,43%

### 5. KESIMPULAN

Berdasarkan hasil analisis dan pembahasan maka diperoleh kesimpulan sebagai berikut:

1. Hasil analisis regresi logistik biner menunjukkan bahwa variabel yang secara serentak mempengaruhi keikutsertaan KB adalah jumlah anak ( $X_3$ ), pendidikan

ayah ( $X_5$ ), dan pendidikan ibu ( $X_6$ ) yang dapat digambarkan dalam model regresi logistik biner sebagai berikut:

$$\pi(x) = \frac{\exp(-1,950 + 1,559 x_3 - 1,600 x_5(1) - 1,210 x_5(2) + 2,086 x_6(1) + 1,611 x_6(1))}{1 + \exp(-1,950 + 1,559 x_3 - 1,600 x_5(1) - 1,210 x_5(2) + 2,086 x_6(1) + 1,611 x_6(1))}$$

2. Model regresi logistik biner yang terbentuk mempunyai nilai ketepatan klasifikasi sebesar 84,375%. Sedangkan hasil analisis FK-NNC menunjukkan bahwa pada  $K = 7$  telah diperoleh nilai ketepatan klasifikasi terbesar yaitu 87,5%.
3. Ketepatan Klasifikasi dengan menggunakan FK-NNC lebih baik daripada klasifikasi Regresi Logistik Biner dalam mengklasifikasikan keikutsertaan KB di kabupaten Klaten tahun 2012.

#### DAFTAR PUSTAKA

- Agresti, A. 2002. *Categorical Data Analysis Second Edition*. John Wiley and Sons. New York.
- Badan Pusat Statistik. 2010. *Hasil Sensus Penduduk 2010 Data Agregat per Provinsi*. BPS. Jakarta.
- Badan Koordinasi Keluarga Berencana Nasional. 2011. *Tata Cara Pelaksanaan Pencatatan Dan Pelaporan Program Kependudukan dan Keluarga Berencana Nasional*. Semarang
- Badan Koordinasi Keluarga Berencana Nasional. *Faktor-Faktor Yang Mempengaruhi Kesertaan Ber-Kb Anggota Kelompok Uppks* [internet]. [www.bkkbn.go.id/litbang/pusna/HasilPenelitian/Faktor-FaktorYangMempengaruhi Kesertaan Ber-Kb Anggota Kelompok Uppks](http://www.bkkbn.go.id/litbang/pusna/HasilPenelitian/Faktor-FaktorYangMempengaruhiKesertaanBer-KbAnggotaKelompokUppks). [Pdf]. (diakses pada tanggal 23 Mei 2014).
- Entjang, I. 1986. *Pendidikan Kependudukan dan Keluarga Berencana*. Penerbit Alumni. Bandung.
- Haloho, O., Pasukat S., Asima M. 2013. Penerapan Analisis Regresi Logistik Pada Pemakaian Alat Kontrasepsi Wanita. *Jurnal Saintia Matematika*. Volume 1 No. 1 (2013), pp. 51–61.
- Hartanto, H. 1994. *Keluarga Berencana dan Kontrasepsi*. Pustaka Sinar Harapan. Jakarta
- Hosmer, D.W., dan Lemenshow. 2000. *Applied Logistic Regression*. USA : John Wiley and Sons.
- Johnson, R. A. dan Wichern, D. W., 1992. *Applied Multivariate Statistical Analysis*. Prentice Hall. New Jersey.
- Prasetyo, E. 2012a. *Data Mining Konsep dan Aplikasi Menggunakan Matlab*. Andi. Yogyakarta.
- Prasetyo. 2012b. *Fuzzy K-Nearest Neighbor in every Class untuk Klasifikasi Data*. *Seminar Nasional Teknik Informatika (SANTIKA 2012)*. pp 57-60.
- Sumini, Yam'ah T., Wahyono K., 2009. *Kontribusi Pemakaian Alat Kontrasepsi Terhadap Fertilitas*. BKKBN.