

## Pengembangan Model Regresi Logistik Multinomial untuk Klasifikasi Politik pada Pemilihan Umum

Vera Maya Santi

*Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta*

Email: vmsanti@unj.ac.id

### Abstract

To fully meet all the requirements of a government 'of the people, by the people, and for the people', politics in a democracy need to make general election. Various political parties in general elections indicate diversity of interest and ideology. The fluctuation number of parties in general election and voter behavior affect fluctuation number of voters. Multinomial logistic regression analysis is one of methods that can be used to classify political parties based on voter behavior. In this study, the response variables based on ideology of electoral parties have three categories: Nationalist-Religious Party (PNR), Nationalist-Secular Party (PNS), and Islamic Party (PI). The explanatory variable is behavior of novice voters. Based on model significance test with a significance level of 0.05, there are 11 variables show significant influence to the classification of political parties in the election simultaneously.

**Keywords:** general election, multinomial logistic regression, ratio odds.

### Abstrak

Untuk memenuhi semua persyaratan pemerintah dari rakyat, oleh rakyat, dan untuk rakyat sepenuhnya, politik demokrasi memerlukan pemilihan umum. Berbagai partai politik dalam pemilihan umum menunjukkan keragaman minat dan ideologi. Jumlah fluktuasi partai dalam pemilihan umum dan perilaku pemilih mempengaruhi fluktuasi jumlah pemilih. Analisis regresi logistik multinomial adalah salah satu metode yang dapat digunakan untuk mengklasifikasikan partai politik berdasarkan perilaku pemilih. Dalam penelitian ini, variabel respon berdasarkan ideologi partai pemilu memiliki tiga kategori: Partai Nasionalis-Agama (PNR), Partai Nasional-Sekuler (PNS), dan Partai Islam (PI). Variabel penjelasnya adalah perilaku pemilih pemula. Berdasarkan uji signifikansi model dengan tingkat signifikansi 0,05, terdapat 11 variabel yang menunjukkan pengaruh signifikan terhadap klasifikasi partai politik dalam pemilu secara bersamaan.

**Kata-kata kunci:** pemilihan umum, regresi logistik multinomial, odd rasio.

### PENDAHULUAN

Pemilihan umum adalah salah satu sarana suksesi politik dalam sebuah negara yang demokratis untuk menghasilkan wakil-wakil rakyat yang mampu mengerti mengenai aspirasi dari rakyat. Disisi lain pemilu juga menentukan partai politik mana yang masih tetap eksis dan paling berhak melanjutkan tugasnya dalam rangka mewujudkan kesejahteraan rakyat.

Partai-partai politik yang mengikuti pemilu memiliki ideologi yang berbeda-beda. Dengan ideologinya masing-masing partai politik akan mempunyai identitas yang jelas, hal itulah kemudian yang memudahkan partai politik tersebut dalam mendapatkan massa pendukung. Di sisi yang lain masyarakat pun akan lebih mudah untuk menentukan partai mana yang sesuai dengan keinginan yang memang memperjuangkan nilai-nilai yang diyakini oleh masyarakat. Berdasarkan ideologi tersebut, partai politik dapat dikelompokkan menjadi tiga kelompok besar, yaitu kelompok Partai Nasionalis Religius (PNR), Partai Nasionalis Sekuler (PNS), dan Partai Islam (PI).

Analisis regresi logistik tepat digunakan untuk membuat model yang menghubungkan jenis partai politik dengan variabel perilaku pemilih setiap daerah provinsi. Hal tersebut dikarenakan variabel respon (Y) memiliki tiga kategori, yaitu kelompok PNR, PNS, dan PI. Sedangkan variabel perilaku pemilih pemula digunakan sebagai variabel penjelas (X). Sehingga lebih tepat untuk menggunakan regresi logistik multinomial (Hosmer dan Lemeshow, 2000, p.31).

Dalam penelitian ini, dibahas tentang pengembangan model regresi logistik multinomial untuk mengklasifikasikan jenis partai politik dengan variabel perilaku pemilih pemula.

## METODE

### Analisis Regresi Logistik

Analisis regresi logistik digunakan dengan tujuan untuk menentukan model yang tepat yang dapat menggambarkan hubungan antara suatu variabel respon dan variabel penjelasnya. Regresi logistik memiliki asumsi-asumsi antara lain adalah kebebasan antar pilihan variabel respon, terdapat hubungan yang linier antara variabel penjelas yang kontinu dan transformasi fungsi logit dari variabel respon, dan kelompok dari variabel respon tidak terklasifikasikan secara sempurna oleh variabel-variabel penjelas (Tabachnick dkk, 2007).

Regresi logistik menggunakan  $\pi(x) = E(Y | x)$  untuk menggambarkan *mean* dari Y dan x. Bentuk parameter yang digunakan untuk model regresi logistik dapat dinyatakan dalam persamaan berikut:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k}} \quad (1)$$

Bentuk penyerhanaan model regresi logistik menggunakan transformasi logit  $\pi(x)$ . Bentuk persamaan logit tersebut adalah sebagai berikut:

$$g(x) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (2)$$

### Analisis Regresi Logistik Multinomial

Analisis regresi logistik multinomial merupakan analisis regresi yang digunakan jika variabel respon bertipe kategorik dengan skala pengukuran nominal. Regresi logistik multinomial merupakan perluasan dari regresi logistik dengan respon biner yang dapat menangani variabel respon dengan lebih dari dua kategori, sehingga mempunyai distribusi multinomial (Montgomery, 2010:9).

Bentuk umum model regresi logistik dapat diformulasikan seperti berikut:

$$E(y) = \pi = \frac{\exp(x' \beta)}{1 + \exp(x' \beta)} \quad (3)$$

Andaikan x sebagai vektor variabel penjelas yang berukuran dengan  $x_0 = 1$ . Maka konstanta dari fungsi logitnya adalah

$$\pi_0(x) = \frac{1}{1 + e^{g_1(x)} + e^{g_2(x)}} \quad (4)$$

dan

$$\pi_j(x) = e^{g_j(x)} \frac{1}{1 + e^{g_1(x)} + e^{g_2(x)}} \quad (5)$$

Untuk  $j = 0, 1, 2$

### Pendugaan Parameter

Variabel respon yang menjadi amatan diasumsikan saling independen, maka diperoleh fungsi log *likelihood* sebagai berikut (Hollander, 2013) :

$$L(\beta) = \sum_{i=1}^n y_{1i} g_1(x_i) + y_{2i} g_2(x_i) - \ln(1 + e^{g_1(x_i)} + e^{g_2(x_i)}) \tag{6}$$

Persamaan *likelihood* dapat diperoleh dengan menurunkan secara parsial sekali dari  $L(\beta)$ . Sedemikian sehingga, menurut teori *maximum likelihood*, untuk menaksirkan varian kovarian dapat diperoleh melalui turunan kedua fungsi *likelihood*.

**Uji Signifikansi Penduga Parameter**

Hipotesis yang digunakan untuk menguji signifikansi parameter dengan Uji Wald adalah sebagai berikut:

$H_0 : \beta_k = 0, k = 1, 2, \dots, k$  (tidak ada pengaruh antara variabel penjelas dengan variabel respon).

$H_1 : \beta_k \neq 0, k = 1, 2, \dots, k$  (ada pengaruh antara variabel penjelas dengan variabel respon).

Statistik uji yang digunakan untuk uji signifikansi adalah (Lemeshow & Hosmer, 2000:16)

$$W_k = \left( \frac{\widehat{\beta}_k}{SE(\widehat{\beta}_k)} \right)^2 \tag{7}$$

Dimana :

$\widehat{\beta}_k$  : Penduga koefisien  $\beta_k$

$SE(\widehat{\beta}_k)$  : Taksiran standar eror parameter

Pengambilan keputusan untuk uji ini yaitu, tolak  $H_0$  jika  $W_k > \chi^2_{1,\alpha}$  atau  $p\text{-value} < \alpha$ . Artinya, variabel penjelas tersebut secara parsial berpengaruh pada variabel respon.

**Uji Kesesuaian Model**

Uji kesesuaian model dilakukan untuk mengetahui adanya perbedaan antara hasil observasi dengan hasil prediksi setelah model serentak terbentuk.

Untuk menguji kesesuaian model dengan hasil obeservasi dapat menggunakan deviansi dengan persamaan sebagai berikut:

$$D = 2 \sum_{i=1}^n \left[ y_i \ln \left( \frac{y_i}{n_i - \pi_i} \right) + (n_i - y_i) \ln \left( \frac{n_i - y_i}{n_i(1 - \pi_i)} \right) \right] \tag{8}$$

D berdistribusi Chi-Square  $\chi^2_{n-p}$  dengan pengambilan keputusan terima  $H_0$  jika  $D > \chi^2_{n-p}$  atau  $Sig > \alpha = 0,05$ . Artinya, tidak ada perbedaan antara hasil observasi dengan hasil prediksi model. (Tabachnick, 2001).

**HASIL DAN PEMBAHASAN**

**Uji Validitas dan Relibilitas**

Data yang digunakan haruslah data valid dan reliabel agar menghasilkan data yang yang dapat dipercaya kebenarannya. Pengujian ini menggunakan analisis pada tiap butir pertanyaan kuesioner dengan mengkorelasikan skor dengan total skor. Dengan statistik uji

$$r = \frac{n \sum_{i=1}^n x_i y_i - \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right)}{\sqrt{\left\{ n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2 \right\} \left\{ n \sum_{i=1}^n y_i^2 - \left( \sum_{i=1}^n y_i \right)^2 \right\}}} \tag{9}$$

didapatkan variabel penjelas agama, suku, tempat tinggal, pendapatan, dan perilaku pemilih, yang terdiri dari citra sosial, identitas partai, citra kandidat, isu dan kebijakan politik, serta peristiwa tertentu

peserta pemilu, yang berpengaruh karena memiliki nilai signifikan  $\leq \alpha = 0,05$  yang sudah valid dan reliabel dengan nilai koefisien  $\alpha_{Cronbach} = 0,645$ .

**Analisis Regresi Multinomial Uji Signifikan**

- Uji Serentak

Uji signifikansi secara serentak dilakukan dengan menggunakan uji rasio *likelihood*. Dengan hasil pengujian  $Sig. = 0,003 < \alpha = 0,05$  dapat diputuskan menolak  $H_0$  yang artinya 10 variabel penjelas secara serentak berpengaruh terhadap karakteristik partai politik.

- Uji Parsial

Sesuai dengan hasil uji signifikansi secara parsial didapatkan fungsi logit yang menjelaskan karakteristik partai politik antara lain sebagai berikut :

$$Y_1 = 2,774X_{8,1} + 1,768X_{9,1} - 24,713X_{10,1} - 24,283X_{11,1} + 4,965X_{12,2} + 8,993X_{13,2} - 1,185X_{14,2} - 0,325X_{15,1} - 18,039X_{16,2} - 0,482X_{17,1} \tag{10}$$

$$Y_2 = 51,459X_{8,2} + 11,521X_{9,2} + 88,205X_{10,2} - 136,608X_{11,1} + 113,045X_{12,2} + 6X_{13,2} - 18,547X_{14,2} + 22,260X_{15,1} - 215,144X_{16,2} - 91,712X_{17,1} \tag{11}$$

Fungsi logit  $Y_1$  dan  $Y_2$  merupakan fungsi yang menunjukkan variabel citra kandidat peserta pemilu yang mempengaruhi secara parsial untuk pemilihan Partai Nasional Religius dan Partai Nasional Sekuler.

**Uji Kesesuaian Model**

Uji kesesuaian model merupakan uji untuk mengetahui perbedaan antara hasil obesrvasi dengan hasil prediksi.

**TABEL 1. GOODNESS OF FIT**

	Chi-Square	Significance
Deviance	41,172	1,000

Berdasarkan tabel diatas, nilai Signifikan = 1,000 > 0,05, artinya terima  $H_0$  yang menunjukkan bahwa tidak ada perbedaan antara hasi observasi dengan hasil prediksi model.

**Interpretasi Model**

Nilai odds untuk fungsi logit PNR sebagai berikut:

1. Kesamaan daerah asal mempunyai nilai rasio odds sebesar 0,215, artinya kecenderungan kesamaan daerah asal memilih PNR sebesar 0,215 kali lipat dibandingkan dengan tidak memiliki kesamaan daerah asal.
2. Kesamaan agama mempunyai nilai rasio odds sebesar 5,859 artinya kecenderungan kesamaan agama memilih PNR sebesar 5,859 kali lipatdibandingkan dengan tidak memiliki kesamaan agama.
3. Latar belakang partai politik mempunyai nilai rasio odds sebesar  $1,851 \times 10^{-11}$  artinya kecenderungan latar belakang partai politik memilih PNR sebesar  $1,851 \times 10^{-11}$  kali lipat kurangnya dibandingkan dengan tidak memilih berdasarkan latar belakangnya.
4. Ideologi partai politik mempunyai nilai rasio odds sebesar  $2,846 \times 10^{-11}$  artinya kecenderungan ideologi partai politik memilih PNR sebesar  $2,846 \times 10^{-11}$  kali lipat kurangnya dibandingkan dengan tidak memilih berdasarkan ideologinya.

5. Informasi peserta pemilu mempunyai nilai rasio odds sebesar 143,323 artinya kecenderungan sedikit mengetahui informasi memilih PNR sebesar 143,323 kali lipat dibandingkan dengan tidak mengetahui informasi peserta.
6. Ketokohan peserta pemilu mempunyai nilai rasio odds sebesar 8045,263 artinya kecenderungan sedikit mempertimbangkan ketokohan memilih PNR sebesar 8045,263 kali lipat dibandingkan dengan tidak mengetahui ketokohan peserta pemilu.
7. Popularitas peserta pemilu mempunyai nilai rasio odds sebesar 0,306 artinya kecenderungan sedikit mementingkan popularitas peserta pemilu memilih PNR sebesar 0,306 kali lipat kurangnya dibandingkan tidak mementingkan popularitas peserta pemilu.
8. Kemampuan Intelektual peserta pemilu mempunyai nilai rasio odds sebesar 0,723 artinya kecenderungan mengetahui kemampuan intelektual peserta pemilu memilih PNR sebesar 0,723 kali lipat kurangnya dibandingkan dengan tidak mengetahui kemampuan intelektual peserta pemilu.
9. Isu dan janji politik mempunyai nilai rasio odds sebesar  $1,465 \times 10^{-8}$  artinya kecenderungan tidak terlalu mempertimbangkan Isu dan janji politik memilih PNR sebesar  $1,465 \times 10^{-8}$  kali lipat dibandingkan mempertimbangkan Isu dan janji politik.
10. Kasus hukum yang menjerat peserta pemilu mempunyai nilai rasio odds sebesar 0,618 artinya kecenderungan memperhatikan kasus hukum yang menjerat peserta pemilu memilih PNR sebesar 0,618 kali lipat kurangnya dibandingkan dengan yang tidak memperhatikan kasus hukum yang menjerat peserta pemilu.

Sedangkan nilai odds untuk fungsi logit Partai Nasionalis Sekuler (PNS) sebagai berikut:

1. Kesamaan daerah asal mempunyai nilai rasio odds sebesar  $2,230 \times 10^{22}$  artinya kecenderungan kesamaan daerah asal sedikit berpengaruh dalam memilih PNS sebesar  $2,230 \times 10^{22}$  kali lipat dibandingkan dengan tidak memiliki kesamaan daerah asal.
2. Kesamaan agama mempunyai nilai rasio odds sebesar 100815,696 artinya kecenderungan kesamaan agama sedikit berpengaruh dalam memilih PNS sebesar 100815,696 kali lipat dibandingkan dengan tidak memiliki kesamaan agama.
3. Latar belakang partai politik mempunyai nilai rasio odds sebesar  $2,027 \times 10^{38}$  artinya kecenderungan latar belakang partai politik memilih PNS sebesar  $2,027 \times 10^{38}$  kali lipat dibandingkan dengan tidak memilih berdasarkan latar belakangnya.
4. Ideologi partai politik mempunyai nilai rasio odds sebesar  $4,698 \times 10^{-60}$  artinya kecenderungan ideologi partai politik memilih PNS sebesar  $4,698 \times 10^{-60}$  kali lipat dibandingkan dengan tidak memilih berdasarkan ideologinya.
5. Informasi peserta pemilu mempunyai nilai rasio odds sebesar  $1,244 \times 10^{49}$  artinya kecenderungan sedikit mengetahui informasi peserta pemilu memilih PNS sebesar  $1,244 \times 10^{49}$  kali lipat dibandingkan dengan yang tidak mengetahui informasi peserta pemilu.
6. Ketokohan peserta pemilu mempunyai nilai rasio odds sebesar 403,619, artinya kecenderungan sedikit mengetahui ketokohan peserta pemilu memilih PNS sebesar 403,619 kali lipat dibandingkan dengan tidak mengetahui ketokohan peserta pemilu.
7. Popularitas peserta pemilu mempunyai nilai rasio odds sebesar  $8,814 \times 10^{-9}$  artinya kecenderungan mementingkan popularitas peserta pemilu memilih PNS sebesar  $8,814 \times 10^{-9}$  kali lipat kurangnya dibandingkan dengan yang tidak mementingkan popularitasnya.
8. Kemampuan Intelektual peserta pemilu mempunyai nilai rasio odds sebesar  $4,651 \times 10^9$  artinya kecenderungan mengetahui kemampuan intelektual peserta pemilu memilih PNS

sebesar  $4,651 \times 10^9$  kali lipat dibandingkan dengan yang tidak mengetahui kemampuan intelektualnya.

9. Isu dan janji politik mempunyai nilai rasio odds sebesar  $3,666 \times 10^{-94}$  artinya kecenderungan tidak terlalu mempertimbangkan Isu dan janji politik memilih PNR sebesar  $3,666 \times 10^{-94}$  kali lipat kurangnya dibandingkan mempertimbangkan Isu dan janji politik.
10. Kasus hukum yang menjerat peserta pemilu mempunyai nilai rasio odds sebesar  $1,479 \times 10^{-40}$  artinya kecenderungan memperhatikan kasus hukum yang menjerat peserta pemilu memilih PNR sebesar  $1,479 \times 10^{-40}$  kali lipat kurangnya dibandingkan dengan tidak memperhatikan kasus-kasus hukum yang menimpa peserta pemilu.

## KESIMPULAN DAN SARAN

### Kesimpulan

Berdasarkan analisis model logistik multinomial, dapat diketahui bahwa 10 variabel berpengaruh secara serentak dengan  $\alpha = 0.05$ . Variabel yang berpengaruh secara parsial terhadap pemilihan Partai Nasionalis Religius dan Partai Nasionalis Sekuler adalah citra sosial, identitas partai, citra kandidat, isu dan kebijakan politik, serta peristiwa tertentu peserta pemilu. Berdasarkan uji kesesuaian model diperoleh signifikansi = 1, yaitu terima  $H_0$ . Artinya tidak ada perbedaan antara hasil observasi dengan hasil prediksi model atau model yang digunakan baik.

### Saran

Setelah diperoleh hasil dengan metode regresi logistik multinomial, penulis menyarankan untuk menggunakan metode lain agar diperoleh hasil yang dapat dibandingkan dengan metode ini. Lalu cari tingkat keakuratan dari model untuk mengetahui seberapa akurat model tersebut.

## UCAPAN TERIMA KASIH

Pada akhir paper ini penulis ingin mengucapkan terimakasih kepada Fakultas Matematika dan Ilmu Pengetahuan Alam serta Program Studi Matematika Universita Negeri Jakarta yang telah memberikan dukungan penuh terhadap penelitian ini serta Ristekdikti yang juga memberikan dukungannya sehingga penelitian ini dapat terselesaikan.

## REFERENSI

- Aminudin, Suryana. 'Perilaku politik di Indonesia'. Jurnal Aspirasi Vol. 1 No. 2 ISSN (online) : 2087-2208. 2011.
- Hollander, *et al.* (2013) *Nonparametric Statistical Methods*. USA: Hohn Wiley & Sons, Inc.
- Hosmer, D. W., Lemeshow, S. (2000) *Applied Logistic Regression 2nd Edition*. New York: John Wiley.
- Menard, S. (2001) *Applied Logistic Regression Analysis Second Edition*. London: Sage Pub.
- Montgomery, *et al.* (2010) *Generalized Linear Models*. USA: John Wiley & Sons, Inc.
- Muhiddin, Amir. (2015). 'Penduduk dan Demokrasi (Studi Kasus pada Pemilihan Gubernur Sulawesi Selatan Tahun 2013-2018 di Kecamatan Somba Opu Kabupaten Gowa'. Jurnal Ilmu Pemerintahan Vol.V No.1. Makassar: Universitas Muhammadiyah Makassar.

- Nurjaman, Asep. (2011). 'Peta Baru Ideologi Partai Politik Indonesia'. Jurnal Ilmiah Bestari. Malang: Universitas Muhammadiyah Malang.
- Tabachnick B.G., Fidell, L.S., Osterlind. (2001)*Using Multivariate Statistics 5th Edition*. USA: Allyn & Bacon Boston.