

PREDIKSI PEMILIH MILENIAL DAN PASCA MILENIAL DI PEMILU MENGUNAKAN ALGORITMA NAIVE BAYES

Oleh:

Sista Faizzah R.

Arif Senja Fitriani

Progam Studi Informatika

Universitas Muhammadiyah Sidoarjo

Mei, 2023

Pendahuluan

- Pemilu adalah peristiwa politik yang penting bagi terwujudnya sistem politik yang demokratis. Ciri negara demokratis ialah negara yang melibatkan masyarakat dalam perencanaan maupun pelaksanaan pemilihan umum
- Masalah utama dalam upaya memprediksi pemilu saat ini terkait adanya fakta bahwa banyaknya milenial yang golput di seluruh wilayah Indonesia, khususnya di Kecamatan Wonokasian, ini dibuktikan dengan tingginya golput antar daerah.
- Penelitian akan diuji berdasarkan data penduduk pemilu yang diperoleh dari Kecamatan Wonokasian tahun 2019 dengan menggunakan metode data mining Naïve Bayes untuk dilakukan analisis untuk memperoleh informasi terhadap data tingkat kehadiran dalam pemilu. Diharapkan dari penelitian yang dilakukan terhadap sampel data penduduk pemilu tersebut dapat diperoleh suatu informasi yang bisa membantu pihak kecamatan untuk merancang strategi dalam meningkatkan kepedulian masyarakat.

Pertanyaan Penelitian (Rumusan Masalah)

- Bagaimana menghitung kehadiran menggunakan weka untuk mengetahui jumlah warga yang golput maupun tidak golput ?

Metode

- $$P(H|X) = \frac{P(X|H) \times P(H)}{P(X)}$$
- X = Data dengan class yang belum diketahui
- H = Hipotesis data X merupakan suatu class spesifik
- $P(H|X)$ = Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob.)
- $P(H)$ = Probabilitas hipotesis H (prior prob.)
- $P(X|H)$ = Probabilitas X berdasarkan kondisi tersebut
- $P(X)$ = Probabilitas dari X

Hasil

- Data Pemilu

No	jml KK	Asal Penduduk	Kategori Masy	Kawin	jenis_kelamin	alamat	rt	Tps	lokasi	hadir kel	hadir
1	Dua	SIDOARJO	Milenial	S	P	WONOKASIAN	rt01	TPS-1	dalam	TSH	TH
2	Dua	SIDOARJO	Milenial	S	L	WONOKASIAN	rt01	TPS-1	dalam	TSH	H
3	Tiga	LUAR	Milenial	B	L	WONOKASIAN	rt01	TPS-1	dalam	THS	TH
4	Tiga	LUAR	PascaMilenial	S	P	WONOKASIAN	rt01	TPS-1	dalam	THS	TH
5	Tiga	LUAR	PascaMilenial	S	L	WONOKASIAN	rt01	TPS-1	dalam	THS	H
6	Dua	SIDOARJO	Milenial	S	L	WONOKASIAN	rt01	TPS-1	dalam	HS	H
7	Dua	SIDOARJO	Milenial	S	P	WONOKASIAN	rt01	TPS-1	dalam	HS	H
8	Dua	SIDOARJO	Milenial	S	L	WONOKASIAN	rt01	TPS-1	dalam	HS	H
9	Dua	SIDOARJO	Milenial	S	P	WONOKASIAN	rt01	TPS-1	dalam	HS	H
10	Dua	SIDOARJO	Milenial	S	L	WONOKASIAN	rt01	TPS-1	dalam	HS	H
..

Hasil

Data training yang digunakan sebanyak 70 data

jml KK	Asal Penduduk	Kategori Masy kawin	jenis_kelamin	alamat	rt	tps	lokasi	hadir kel	hadir	
tiga	SIDOARJO	PascaMilenial	Sudah	Laki-Laki	WONOKASIAN	rt01	TPS-1	dalam	Tidak Hadir Semua	Tidak Hadir
tiga	SIDOARJO	Milenial	Sudah	Perempuan	WONOKASIAN	rt01	TPS-1	dalam	Tidak Hadir Semua	Tidak Hadir
tiga	SIDOARJO	Pemula	Belum	Perempuan	WONOKASIAN	rt01	TPS-1	dalam	Tidak Hadir Semua	Tidak Hadir
dua	SIDOARJO	Milenial	Sudah	Perempuan	WONOKASIAN	rt01	TPS-1	dalam	Hadir Semua	Hadir
dua	SIDOARJO	PascaMilenial	Sudah	Laki-Laki	WONOKASIAN	rt01	TPS-1	dalam	Hadir Semua	Hadir
tiga	SIDOARJO	Milenial	Belum	Laki-Laki	WONOKASIAN	rt01	TPS-1	dalam	Hadir Semua	Hadir
tiga	SIDOARJO	Milenial	Sudah	Perempuan	WONOKASIAN	rt01	TPS-1	dalam	Hadir Semua	Hadir
tiga	SIDOARJO	PascaMilenial	Sudah	Laki-Laki	WONOKASIAN	rt01	TPS-1	dalam	Hadir Semua	Hadir
dua	SIDOARJO	PascaMilenial	Sudah	Perempuan	WONOKASIAN	rt01	TPS-1	dalam	Tidak Semua Hadir	Hadir
dua	SIDOARJO	PascaMilenial	Sudah	Laki-Laki	WONOKASIAN	rt01	TPS-1	dalam	Tidak Semua Hadir	Tidak Hadir
..

Hasil

Data testing yang digunakan sebanyak 30 data

jml KK	Asal Penduduk	Kategori Masy	kawin	jenis_kelamin	alamat	rt	tps	lokasi	hadir kel	hadir
tiga	SIDOARJO	PascaMilennial	Sudah	Laki-Laki	WONOKASIAN	rt01	TPS-1	dalam	Tidak Hadir Semua	?
tiga	SIDOARJO	Milenial	Sudah	Perempuan	WONOKASIAN	rt01	TPS-1	dalam	Tidak Hadir Semua	?
tiga	SIDOARJO	Pemula	Belum	Perempuan	WONOKASIAN	rt01	TPS-1	dalam	Tidak Hadir Semua	?
dua	SIDOARJO	Milenial	Sudah	Perempuan	WONOKASIAN	rt01	TPS-1	dalam	Hadir Semua	?
dua	SIDOARJO	PascaMilennial	Sudah	Laki-Laki	WONOKASIAN	rt01	TPS-1	dalam	Hadir Semua	?
tiga	SIDOARJO	Milenial	Belum	Laki-Laki	WONOKASIAN	rt01	TPS-1	dalam	Hadir Semua	?
tiga	SIDOARJO	Milenial	Sudah	Perempuan	WONOKASIAN	rt01	TPS-1	dalam	Hadir Semua	?
tiga	SIDOARJO	PascaMilennial	Sudah	Laki-Laki	WONOKASIAN	rt01	TPS-1	dalam	Hadir Semua	?
dua	SIDOARJO	PascaMilennial	Sudah	Perempuan	WONOKASIAN	rt01	TPS-1	dalam	Tidak Semua Hadir	?
dua	SIDOARJO	PascaMilennial	Sudah	Laki-Laki	WONOKASIAN	rt01	TPS-1	dalam	Tidak Semua Hadir	?
..

Pembahasan

Hasil klasifikasi 100 dataset menggunakan weka

Dari hasil pengujian diperoleh tingkat kebenaran sebesar 76,66% pada *Correctly Classified Instances* dan tingkat kesalahan sebesar 23,33% pada *Incorrectly Classified Instances*.

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier: Choose NaiveBayes

Test options

- Use training set
- Supplied test set Set...
- Cross-validation Folds 10
- Percentage split % 66

More options...

(Nom) hadir kel

Start Stop

Result list (right-click for options)

- 15:32:57 - bayes.NaiveBayes

Classifier output

[total] 17.0 5.0 14.0

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	23	76.6667 %
Incorrectly Classified Instances	7	23.3333 %
Kappa statistic	0.593	
Mean absolute error	0.1857	
Root mean squared error	0.3397	
Relative absolute error	46.7641 %	
Root relative squared error	76.2673 %	
Total Number of Instances	30	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.733	0.200	0.786	0.733	0.759	0.535	0.847	0.787	Tidak Semua Hadir
	0.000	0.074	0.000	0.000	0.000	-0.089	0.784	0.239	Tidak Hadir Semua
	1.000	0.111	0.857	1.000	0.923	0.873	1.000	1.000	Hadir Semua
Weighted Avg.	0.767	0.152	0.736	0.767	0.749	0.608	0.902	0.817	

=== Confusion Matrix ===

```
a b c <-- classified as
11 2 2 | a = Tidak Semua Hadir
3 0 0 | b = Tidak Hadir Semua
0 0 12 | c = Hadir Semua
```

Status OK

Log x 0

Pembahasan

Hasil klasifikasi 100% dari 100 dataset menggunakan weka

Dari hasil pengujian diperoleh tingkat kebenaran sebesar 100% pada *Correctly Classified Instances* dan tingkat kesalahan sebesar 0% pada *Incorrect Classified Instances*.

The screenshot shows the Weka Classifier interface with the NaiveBayes classifier selected. The 'Test options' section is set to 'Use training set'. The 'Classifier output' section displays the following results:

```
=== Evaluation on training set ===
Time taken to test model on training data: 0 seconds

=== Summary ===
Correctly Classified Instances      5      100 %
Incorrectly Classified Instances    0       0 %
Kappa statistic                    1
Mean absolute error                 0.0051
Root mean squared error             0.0078
Relative absolute error             1.0448 %
Root relative squared error         1.5987 %
Total Number of Instances          5

=== Detailed Accuracy By Class ===
          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
          1.000    0.000    1.000     1.000    1.000     1.000  1.000    1.000    Hadir
          1.000    0.000    1.000     1.000    1.000     1.000  1.000    1.000    Tidak Hadir
Weighted Avg.  1.000    0.000    1.000     1.000    1.000     1.000  1.000    1.000

=== Confusion Matrix ===
 a b  <-- classified as
 3 0 | a = Hadir
 0 2 | b = Tidak Hadir
```

Pembahasan

Hasil persentase 100 dataset

Dari hasil klasifikasi menggunakan metode Naïve Bayes untuk mengklasifikasi status kehadiran pada 100 dataset, diperoleh hasil tertinggi menggunakan perhitungan manual dengan perolehan persentase sebesar 80%.

	Dataset	Training	Testing	Benar	% Benar	Salah	% Salah
Weka	100	70	30	23	76,67%	7	23,33%
Manual	100	70	30	24	80%	6	20%

Pembahasan

Hasil persentase 10% dari 100 dataset

Dari hasil klasifikasi menggunakan metode Naïve Bayes untuk mengklasifikasi status kehadiran pada 10% dari 100 dataset, diperoleh hasil tertinggi menggunakan perhitungan weka dengan perolehan persentase sebesar 100%.

Perbandingan	Dataset	Training	Testing	Benar	% Benar	Salah	% Salah
Manual	100	5	5	2	60%	3	40%
Weka	100	5	5	5	100%	0	0%
Weka	4.249	-	-	3.846	90,51%	403	9,48%

Manfaat Penelitian

Manfaat dari hasil penelitian ini diharapkan dapat memprediksi hadir dan tidaknya warga saat pemilu berlangsung sehingga memudahkan panitia untuk mengetahui.

Referensi

- W.-K. Chen, *Linear Networks and Systems*. Belmont, CA: Wadsworth, 1993, pp. 123-135.
- R. Hayes, G. Pisano, D. Upton, and S. Wheelwright, *Operations, Strategy, and Technology: Pursuing the competitive edge*. Hoboken, NJ: Wiley, 2005.
- The Oxford Dictionary of Computing, 5th ed. Oxford: Oxford University Press, 2003.
- A. Rezi and M. Allam, "Techniques in array processing by means of transformations," in *Control and Dynamic Systems*, Vol. 69, Multidimensional Systems, C. T. Leondes, Ed. San Diego: Academic Press, 1995, pp. 133-180.
- O. B. R. Strimpel, "Computer graphics," in *McGraw-Hill Encyclopedia of Science and Technology*, 8th ed., Vol. 4. New York: McGraw-Hill, 1997, pp. 279-283.
- H. Ayasso and A. Mohammad-Djafari, "Joint NDT Image Restoration and Segmentation Using Gauss–Markov–Potts Prior Models and Variational Bayesian Computation," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2265-77, 2010. [Online]. Available: IEEE Xplore, <http://www.ieee.org>. [Accessed Sept. 10, 2010].
- A. Altun, "Understanding hypertext in the context of reading on the web: Language learners' experience," *Current Issues in Education*, vol. 6, no. 12, July 2003. [Online]. Available: <http://cie.ed.asu.edu/volume6/number12/>. [Accessed Dec. 2, 2004].
- H. Imron, R. R. Isnanto and E. D. Widiyanto, "Perancangan Sistem Kendali pada Alat Listrik Rumah Tangga Menggunakan Media Pesan Singkat (SMS)". *Jurnal Teknologi dan Sistem Komputer*, vol.4, no. 3, pp. 454-462, Agustus 2016. [Online]. doi: <http://dx.doi.org/10.14710/4.3.2016.454-462>. [Diakses 4 September 2016].
- J. R. Beveridge and E. M. Riseman, "How easy is matching 2D line models using local search?" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 564-579, June 1997.
- E. H. Miller, "A note on reflector arrays," *IEEE Transactions on Antennas and Propagation*, to be published.

