

ANALISIS KEHDIRAN PEMILIH MENGUNAKAN ALGORITMA C4.5 PADA SISTEM PEMILU

Oleh:

Wahidiyah Kurniawati,

Arif Senja Fitriani

Progam Studi Informatika

Universitas Muhammadiyah Sidoarjo

Juli, 2023

Pendahuluan

Ciri negara demokratis ialah negara yang melibatkan masyarakat dalam perencanaan maupun pelaksanaan pemilihan umum. Partisipasi masyarakat (pemilih) merupakan aspek penting dalam tatanan negara demokrasi. Misalnya dalam Pemilu partisipasi politik berpengaruh terhadap legitimasi masyarakat kepada calon atau pasangan calon yang terpilih. Setiap masyarakat memiliki preferensi dan kepentingan masing-masing untuk menentukan pilihan mereka dalam pemilu.

Meningkatnya jumlah pemilih yang mengikuti pemilu dapat dijadikan sebagai patokan keberhasilan pemilu. Kebijakan voting pada pemilu dapat ditentukan dengan beberapa metode untuk mendapatkan nilai dari sekumpulan ide dan data dengan menggunakan ilmu data mining.

Berdasarkan pemikiran tersebut, maka peneliti melakukan penelitian yang membahas tentang “Prediksi Partisipasi Pada Pemilu 2019 Desa Wonokasian Menggunakan Algoritma C4.5”.

Pertanyaan Penelitian (Rumusan Masalah)

- Bagaimana cara memprediksi partisipasi pemilih dengan menggunakan algoritma C4.5?

Metode

- **Metode C4.5**

$$Entropy = \sum_{i=1}^n -p_i * \log_2 p_i$$

Keterangan:

S: Himpunan kasus

A: Atribut

N: Jumlah partisi S

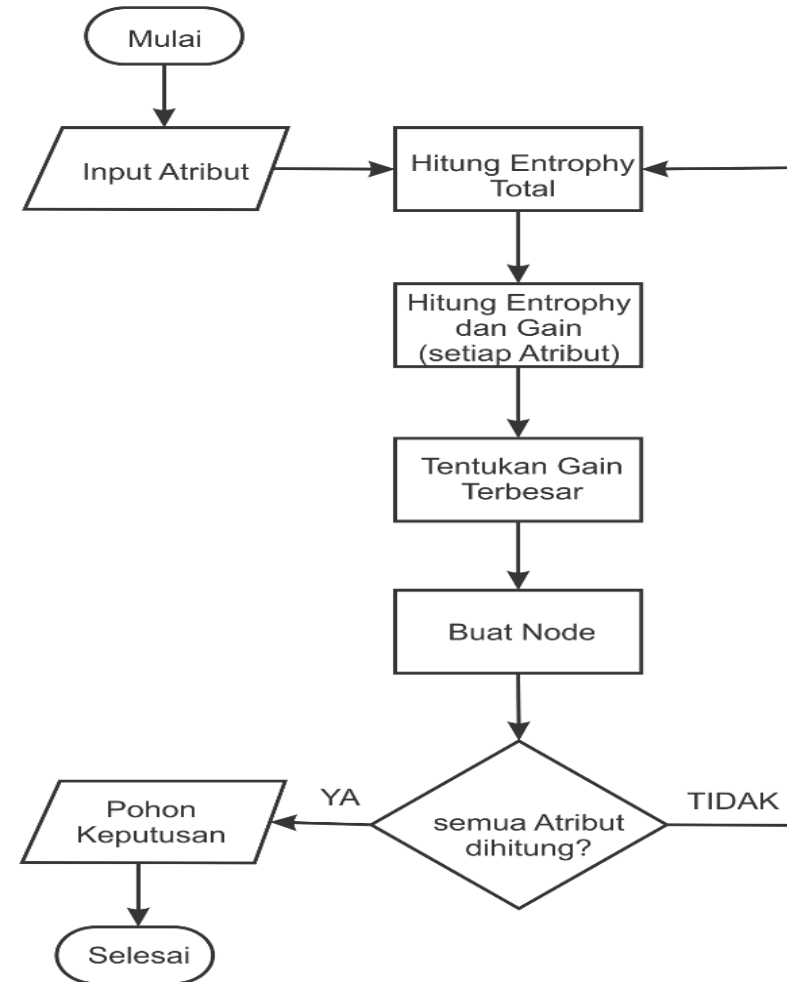
P_i : Proporsi dari S_i terhadap S

Nialah jumlah nilai yang terdapat pada atribut target (jumlah kelas). Sedangkan P_i menyatakan porsi / rasio antara jumlah sampel di kelas i dengan jumlah sampel pada himpunan data.

Hasil

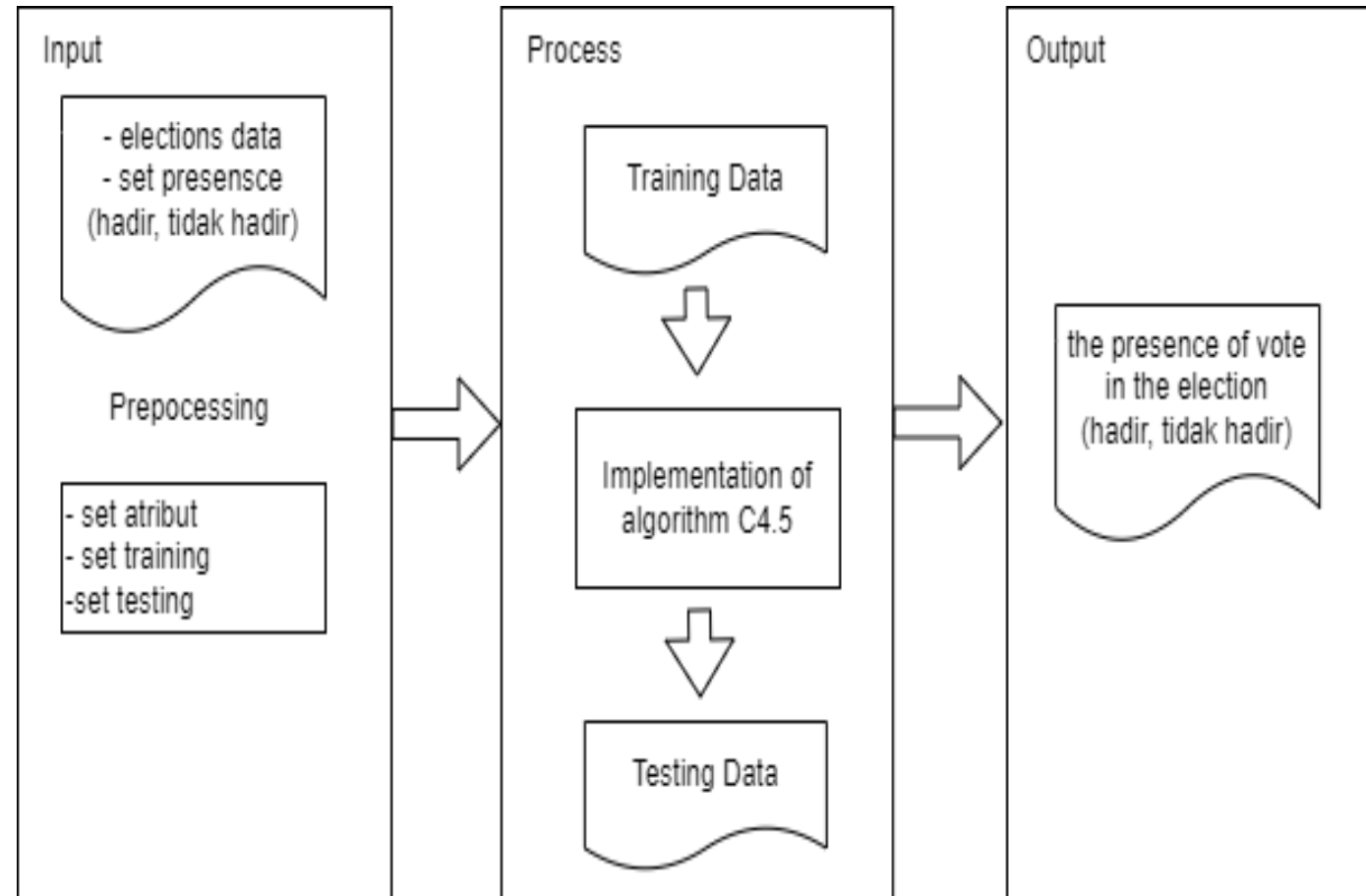
- **Flowchart**

Secara umum, pengertian flowchart merupakan gambaran dari langkah-langkah dan urutan prosedur sebuah program secara grafik. Flowchart digunakan untuk analisis dan programmer dalam menyelesaikan masalah dengan membaginya ke dalam segmen-segmen yang lebih kecil dan digunakan untuk menganalisis alternatif-alternatif lain dalam pengoperasian.



Hasil

- **Diagram Penelitian**



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
..

Pembahasan

•Data Training

Split Training Set 60% (2549 instances)

No	jml KK	Asal Penduduk	Kategori Masy	kawin	jenis_kelamin	rt	tps	lokasi	hadir kel	hadir
1	dua	SIDOARJO	Milenial	S	P	rt01	TPS-1	dalam	TSH	TH
2	dua	SIDOARJO	Milenial	S	L	rt01	TPS-1	dalam	TSH	H
3	tiga	LUAR	Milenial	B	L	rt01	TPS-1	dalam	THS	TH
4	tiga	LUAR	PascaMilenial	S	P	rt01	TPS-1	dalam	THS	TH
5	tiga	LUAR	PascaMilenial	S	L	rt01	TPS-1	dalam	THS	H
6	dua	SIDOARJO	Milenial	S	L	rt01	TPS-1	dalam	HS	H
7	dua	SIDOARJO	Milenial	S	P	rt01	TPS-1	dalam	HS	H
8	dua	SIDOARJO	Milenial	S	L	rt01	TPS-1	dalam	HS	H
9	dua	SIDOARJO	Milenial	S	P	rt01	TPS-1	dalam	HS	H
10	dua	SIDOARJO	Milenial	S	L	rt01	TPS-1	dalam	HS	H
..
2545	empat	SIDOARJO	Milenial	B	L	rt12	TPS-10	luar	HS	H
2546	empat	SIDOARJO	PascaMilenial	S	P	rt12	TPS-10	luar	HS	H
2547	empat	SIDOARJO	Milenial	B	L	rt12	TPS-10	luar	HS	H
2548	empat	SIDOARJO	PascaMilenial	S	L	rt12	TPS-10	luar	HS	H
2549	tiga	SIDOARJO	Milenial	B	L	rt12	TPS-10	luar	HS	H

Pembahasan

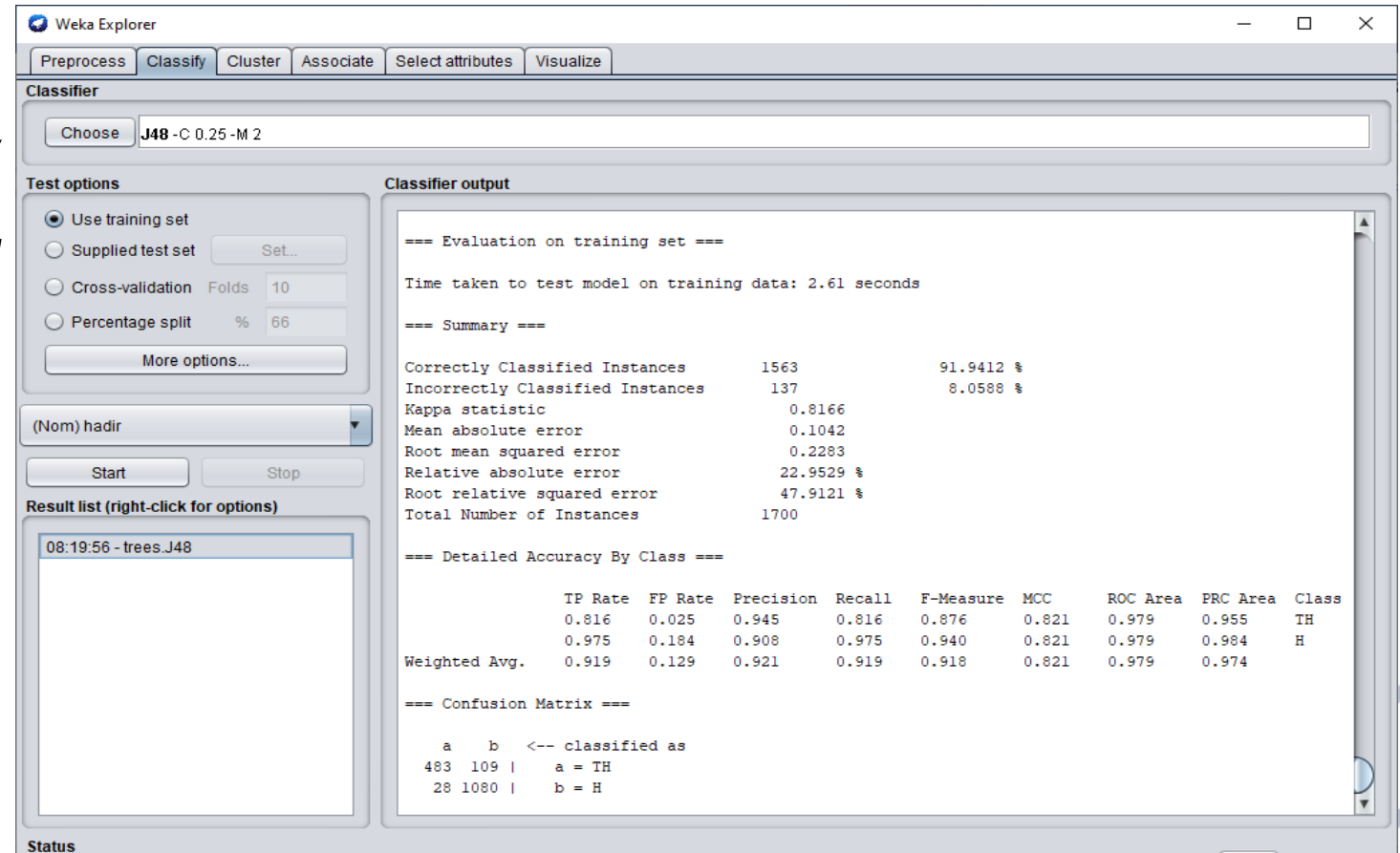
•Data Testing

Split Testing Set 40% (1.700 instances)

No	jml KK	Asal Penduduk	Kategori Masy	kawin	jenis_kelamin	rt	tps	lokasi	hadir kel	hadir
1	tiga	SIDOARJO	PascaMilennial	S	P	rt12	TPS-10	luar	HS	H
2	tiga	SIDOARJO	PascaMilennial	S	L	rt12	TPS-10	luar	HS	H
3	satu	SIDOARJO	Pemula	S	L	rt12	TPS-10	luar	HS	H
4	lima	SIDOARJO	Milenial	B	L	rt12	TPS-10	luar	HS	H
5	lima	SIDOARJO	Milenial	B	L	rt12	TPS-10	luar	HS	H
6	lima	SIDOARJO	PascaMilennial	S	L	rt12	TPS-10	luar	HS	H
7	lima	SIDOARJO	Milenial	B	P	rt12	TPS-10	luar	HS	H
8	lima	SIDOARJO	PascaMilennial	S	P	rt12	TPS-10	luar	HS	H
9	empat	SIDOARJO	Milenial	B	L	rt12	TPS-10	luar	HS	H
10	empat	SIDOARJO	PascaMilennial	S	L	rt12	TPS-10	luar	HS	H
..
1696	dua	LUAR PROV	PascaMilennial	S	L	rt21	TPS-16	dalam	THS	TH
1697	dua	SIDOARJO	PascaMilennial	S	L	rt21	TPS-16	dalam	HS	H
1698	dua	SIDOARJO	PascaMilennial	S	P	rt21	TPS-16	dalam	HS	H
1699	dua	SIDOARJO	Milenial	B	P	rt21	TPS-16	dalam	HS	H
1700	dua	SIDOARJO	Milenial	B	L	rt21	TPS-16	dalam	HS	H

Pembahasan

- Hasil Klasifikasi Menggunakan Weka
Dari hasil pengujian diperoleh tingkat kebenaran sebesar 91,94% pada *Correctly Classified Instances* dan tingkat kesalahan sebesar 8,05% pada *Incorrect Classified Instances*.



Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier: Choose J48 -C 0.25 -M 2

Test options

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

More options...

(Nom) hadir

Start Stop

Result list (right-click for options)

08:19:56 - trees.J48

Status

Classifier output

```
=== Evaluation on training set ===
Time taken to test model on training data: 2.61 seconds

=== Summary ===
Correctly Classified Instances      1563      91.9412 %
Incorrectly Classified Instances    137       8.0588 %
Kappa statistic                    0.8166
Mean absolute error                 0.1042
Root mean squared error             0.2283
Relative absolute error             22.9529 %
Root relative squared error         47.9121 %
Total Number of Instances          1700

=== Detailed Accuracy By Class ===
                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                0.816   0.025   0.945     0.816   0.876     0.821   0.979    0.955    TH
                0.975   0.184   0.908     0.975   0.940     0.821   0.979    0.984    H
Weighted Avg.   0.919   0.129   0.921     0.919   0.918     0.821   0.979    0.974

=== Confusion Matrix ===
  a  b  <-- classified as
483 109 |  a = TH
 28 1080 |  b = H
```

Manfaat Penelitian

- Penulis :
 1. Salah satu cara untuk mengaplikasikan ilmu yang diperoleh selama perkuliahan.
 2. Menambah pengetahuan dalam bidang mengumpulkan data khususnya algoritma C4.5.
 3. Penelitian ini sebagai salah satu syarat untuk memperoleh gelar sajana (Strata-1) program studi Informatika di Universitas Muhammadiyah Sidoarjo.
- Pembaca :
 1. Dapat digunakan sebagai tambahan pengetahuan dalam penerapan data mining untuk klasifikasi prediksi tingkat partisipasi pemilih menggunakan algoritma C4.5.

Referensi

- [1] D. M. Liando, "J. LPPM Bid. EkoSosBudKum," "Pemilu dan Partisipasi Politik Masyarakat (Studi Pada Pemilihan Anggota Legislatif Dan Pemilihan Presiden Dan Calon Wakil Presiden Di Kabupaten Minahasa Tahun 2014)", Vol. 1 dari 2 vol. 3, no. 2, pp. 14–28, 2016.
- [2] A. S. F. Dkk, "International Conference on Engineering, Technologies, and Applied Sciences (ICETsAS)," "Classification Using C4.5 Algorithm in Election Participation Prediction", 2019.
- [3] E. d. H. Amalia, "Algoritma C4.5 Untuk Prediksi Hasil Pemilihan Legislatif DPRD DKI Jakarta" Techno Nusa Mandiri, vol. Vol. IX No.1, 2013.

