

Nama : Wahyu Adi Nugroho

NIM : A11.2019.12310

Kelp. : 46UG

TUGAS DATA MINING

PERTEMUAN MINGGU KE – 2

1. Carilah data yang dapat Anda gunakan untuk Proses Data Mining kemudian gunakan CRISP DM untuk menyelesaikan masalah tersebut !
2. Jelaskan setiap langkahnya seperti Studi Kasus Data Kelulusan Mahasiswa di atas menggunakan proses CRISP DM dimulai dari Business Understanding Data Understanding Data Preparation Modeling Evaluation dan Deployment !

Jawaban :

1. Data yang digunakan untuk proses data mining pada kasus ini adalah data pemilihan umum KPU yang berekstensi .xls.
2. Proses CRISP DM :

- **Pemahaman Proses Bisnis / Bussiness Understanding**

- a. Problem

- Mbak PM adalah seorang anggota / kader dari partai P dari daerah tertentu.
- Sebagai seorang kader, tentunya terjun di dalam dunia politik baik legislatif maupun eksekutif merupakan keinginannya.
- Pada pemilihan umum (pemilu) legislatif 2024 mendatang, Mbak PM ingin mencalonkan dirinya sebagai calon legislatif (caleg).
- Namun, Mbak PM memiliki masalah karena dirinya merasa tidak yakin bisa memenangkan atau memprediksi tingkat keterpilihan dirinya di dalam pemilu tersebut.
- Oleh karena itu, Mbak PM ingin memahami dan membuat pola dari profil caleg sebelumnya yang bisa terpilih dan tidak menjadi anggota legislatif di daerahnya.
- Dengan menggunakan pola tersebut, Mbak PM dapat melakukan prediksi tingkat keterpilihannya sehingga dia tidak ragu lagi membuat keputusan untuk tetap mencalonkan sebagai caleg dan pasti terpilih jika tingkat keterpilihannya tinggi atau mengurungkan niatnya agar terhindar dari kerugian materi imbas dari kampanye pemilu jika tetap nekat mencalonkan diri sedangkan peluang untuk memenangkan pemilu di daerahnya tipis.

- b. Objektif

Menemukan pola dari calon legislatif yang terpilih dan tidak.

Nama : Wahyu Adi Nugroho

NIM : A11.2019.12310

Kelp. : 46UG

- **Pemahaman Data / Data Understanding**

Untuk meyelesaikan masalah tersebut, Mbak PM mengambil dari dari website Situng KPU. Data – data dikumpulkan dari data pemilihan umum kpu dengan atribut seperti berikut.

- NAMA PARPOL.
- NAMA CALON LEGISLATIF.
- JENIS KELAMIN : L untuk laki – laki dan P untuk Perempuan.
- KECAMATAN.
- NO. URUT PARPOL.
- SUARA SAH PARTAI.
- JUML. PEROLEHAN KURSI.
- DAERAH PEMILIHAN.
- NO. URUT CALEG.
- SUARA SAH CALEG.
- STATUS KETERPILIHAN : Terpilih atau Tidak.

- **Persiapan Data / Data Preparation**

- Dataset : datapemilukpu.xls

	B	C	D	E	F	G	H	I	J	K
	NAMA CALON LEGISLATIF	JENIS KELAMIN	KECAMATAN	NO.URUT PARPOL	SUARA SAH PARTAI	JUML.PEROLEHAN KURSI	DAERAH PEMILIHAN	NO.URUT CALEG	SUARA SAH CALEG	TERPILIH ATAU TIDAK
1	TOTO SUKISNO BSC	L	LEBAKSU	1	18578	1	1	1	594	TIDAK
2	EDI PURYANTO SH	L	SLAWI	1	18578	1	1	2	943	TIDAK
3	ELU RETNOWATILSH	P	SLAWI	1	18578	1	1	3	1730	TIDAK
4	SAHYUDIN	L	DUKUHWARU	1	18578	1	1	4	2508	YA
5	H.FAJAR SIGIT KUSUMAJAYA SH	L	SLAWI	1	18578	1	2	1	923	TIDAK
6	SUMIRAH	P	TARUB	1	18578	1	2	2	308	TIDAK
7	DARYOTO	L	TARUB	1	18578	1	2	3	54	TIDAK
8	KHAPP ARRONI S Pdi	L	BOJONG	1	18578	1	3	1	1682	TIDAK
9	ENDANG SUICI RAHAYU	P	JATINEGARA	1	18578	1	3	2	918	TIDAK
10	KH CHAFIDZ SA MUFTI LC	L	SLAWI	1	18578	1	3	3	87	TIDAK
11	H.WALUYO SP	L	BALAPULANG	1	18578	1	4	1	728	TIDAK
12	DARUITO SE	L	MARGASARI	1	18578	1	4	2	346	TIDAK
13	DIWIINDAH LISTYANNINGRUM	P	LEBAKSU	1	18578	1	4	3	184	TIDAK
14	SUSANTO	L	LEBAKSU	1	18578	1	5	1	381	TIDAK
15	ALMI	L	ADWIERNA	1	18578	1	5	2	148	TIDAK
16	POERWO LESTARI	P	SLAWI	1	18578	1	5	3	128	TIDAK
17	KUSMORO	L	SLAWI	1	18578	1	6	1	441	TIDAK
18	NANANG JUWAHR SH	L	BOJONG	1	18578	1	6	2	88	TIDAK
19	USWATUN KHASANAH	P	BALAPULANG	1	18578	1	6	3	74	TIDAK
20	SUPRIYATNI	P	PAGERBARANG	2	12373	0	1	1	351	TIDAK
21	MULIATI FARDILAH	P	PANGAH	2	12373	0	2	1	972	TIDAK
22	TANURI	L	WARUREJA	2	12373	0	2	2	307	TIDAK
23	SROD BSC	L	MARGASARI	2	12373	0	3	1	274	TIDAK
24	KASMIJI	L	MARGASARI	2	12373	0	3	2	186	TIDAK
25	MUSTOFA	L	MARGASARI	2	12373	0	4	1	844	TIDAK
26	ABU NAWAS	L	MARGASARI	2	12373	0	4	2	333	TIDAK
27	MASANI SUKIM SH	L	ADWIERNA	2	12373	0	5	1	804	TIDAK
28	NIURCHIKMAH	P	ADWIERNA	2	12373	0	5	2	275	TIDAK
29	NANING SUDIARNI	L	WARUREJA	2	12373	0	6	1	848	TIDAK
30	KAMARI	L	KRAMAT	2	12373	0	6	2	461	TIDAK
31	MUJAHIDIN	L	SURODADI	2	12373	0	6	3	119	TIDAK

Nama : Wahyu Adi Nugroho

NIM : A11.2019.12310

Kelp. : 46UG

Row No.	TERPILIH AT...	NAMA PART...	NAMA CALO...	JENIS KELA...	KECAMATAN	NO.URUT PA...	SUARA SAH ...	JUML.PERO...	DAERAH PE...
1	TIDAK	HATI NURANI...	TOTO SUKIS...	L	LEBAKSIU	1	18578	1	1
2	TIDAK	HATI NURANI...	EDI PURYAN...	L	SLAWI	1	18578	1	1
3	TIDAK	HATI NURANI...	ELI RETNOW...	P	SLAWI	1	18578	1	1
4	YA	HATI NURANI...	SAHYUDIN	L	DUKUHWARU	1	18578	1	1
5	TIDAK	HATI NURANI...	H.FAJAR SIGI...	L	SLAWI	1	18578	1	2
6	TIDAK	HATI NURANI...	SUMIRAH	P	TARUB	1	18578	1	2
7	TIDAK	HATI NURANI...	DARYOTO	L	TARUB	1	18578	1	2
8	TIDAK	HATI NURANI...	KHAPIP APR...	L	BOJONG	1	18578	1	3
9	TIDAK	HATI NURANI...	ENDANG SU...	P	JATINEGARA	1	18578	1	3
10	TIDAK	HATI NURANI...	KH.CHAFIDZ ...	L	SLAWI	1	18578	1	3
11	TIDAK	HATI NURANI...	H.WALUYO,SIP	L	BALAPULANG	1	18578	1	4
12	TIDAK	HATI NURANI...	DARJITO,SE	L	MARGASARI	1	18578	1	4
13	TIDAK	HATI NURANI...	DWI INDAH L...	P	LEBAKSIU	1	18578	1	4
14	TIDAK	HATI NURANI...	SUSANTO	L	LEBAKSIU	1	18578	1	5

- b. Terdapat 425 data caleg dengan 11 atribut (1 special attributes/label, 10 regular attributes) dengan missing value sebanyak 6 buah dan tidak terdapat data noise.

Name	Type	Missing	Statistics	Filter (11 / 11 attributes):
NO.URUT CALEG	Integer	5	Min 1, Max 11, Average 3.362	
JENIS KELAMIN	Binominal	1	Negative L, Positive P, Values L (306), P (118)	
TERPILIH ATAU TIDAK	Binominal	0	Negative TIDAK, Positive YA, Values TIDAK (377), YA (48)	
NAMA PARTAI POLITIK	Nominal	0	Negative PARTAI PELOPOR (1), Positive PARTAI D [...] NGAN (56), PARTAI D [...] ERJUANGAN (56),	
NAMA CALON LEGESLATIF	Nominal	0	Least dr MOHAM [...] ARIFI (1), Most ROKHANI (3), Values ROKHANI (3), KASMUI (2), ...[416	
KECAMATAN	Nominal	0	Least PANGAH (1), Most SLAWI (64), Values SLAWI (64), KRAMAT (45), ...[35	
NO.URUT PARPOL	Integer	0	Min 1, Max 44, Average 18.073	

Showing attributes 1 - 11 Examples: 425 Special Attributes: 1 Regular Attributes: 10

- c. Missing Value dipecahkan dengan menambahkan data dengan nilai rata rata. • Hasilnya adalah data bersih tanpa missing value.

Name	Type	Missing	Statistics	Filter (11 / 11 attributes):
JENIS KELAMIN	Binominal	0	Negative L, Positive P, Values L (307), P (118)	
KECAMATAN	Polynomial	0	Least PANGAH (1), Most SLAWI (64), Values SLAWI (64), KRAMAT (4)	
NO.URUT PARPOL	Integer	0	Min 1, Max 44, Average 18.073	
SUARA SAH PARTAI	Integer	0	Min 56, Max 143817, Average 52636.075	
JUML.PEROLEHAN KURSI	Integer	0	Min 0, Max 13, Average 4.605	
DAERAH PEMILIHAN	Integer	0	Min 1, Max 6, Average 3.565	
NO.URUT CALEG	Integer	0	Min 1, Max 11, Average 3.358	

Showing attributes 1 - 11 Examples: 425 Special Attributes: 1 Regular Attributes: 10

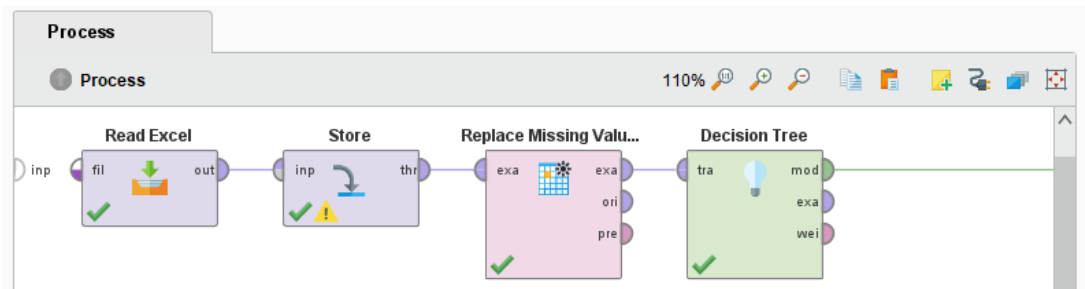
Nama : Wahyu Adi Nugroho

NIM : A11.2019.12310

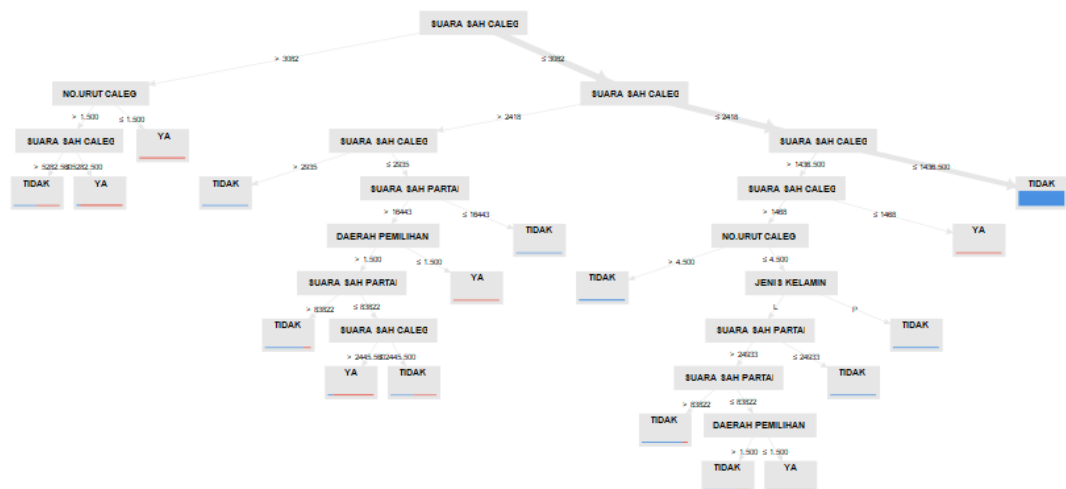
Kelp. : 46UG

- **Pemodelan / Modeling**

Memodelkan dengan Decision Tree.



Hasil pola dari data berupa berupa decision tree pohon keputusan sebagai berikut.



- **Evaluasi / Evaluation**

```
SUARA SAH CALEG > 3082
| NO.URUT CALEG > 1.500
| | SUARA SAH CALEG > 5282.500: TIDAK {TIDAK=2, YA=2}
| | SUARA SAH CALEG ≤ 5282.500: YA {TIDAK=1, YA=11}
| NO.URUT CALEG ≤ 1.500: YA {TIDAK=0, YA=10}
SUARA SAH CALEG ≤ 3082
| SUARA SAH CALEG > 2418
| | SUARA SAH CALEG > 2935: TIDAK {TIDAK=5, YA=0}
| | SUARA SAH CALEG ≤ 2935
| | | SUARA SAH PARTAI > 16443
| | | | DAERAH PEMILIHAN > 1.500
| | | | | SUARA SAH PARTAI > 83822: TIDAK {TIDAK=5, YA=1}
| | | | | SUARA SAH PARTAI ≤ 83822
| | | | | | SUARA SAH CALEG > 2445.500: YA {TIDAK=1, YA=7}
| | | | | | SUARA SAH CALEG ≤ 2445.500: TIDAK {TIDAK=1, YA=1}
| | | | | | DAERAH PEMILIHAN ≤ 1.500: YA {TIDAK=0, YA=4}
| | | | SUARA SAH PARTAI ≤ 16443: TIDAK {TIDAK=2, YA=0}
| SUARA SAH CALEG ≤ 2418
| | SUARA SAH CALEG > 1436.500
| | | SUARA SAH CALEG > 1468
| | | | NO.URUT CALEG > 4.500: TIDAK {TIDAK=15, YA=0}
| | | | NO.URUT CALEG ≤ 4.500
| | | | | JENIS KELAMIN = L
| | | | | | SUARA SAH PARTAI > 24933
| | | | | | | SUARA SAH PARTAI > 83822: TIDAK {TIDAK=8, YA=1}
| | | | | | | SUARA SAH PARTAI ≤ 83822
| | | | | | | | DAERAH PEMILIHAN > 1.500: TIDAK {TIDAK=8, YA=5}
| | | | | | | | DAERAH PEMILIHAN ≤ 1.500: YA {TIDAK=1, YA=3}
| | | | | | | SUARA SAH PARTAI ≤ 24933: TIDAK {TIDAK=6, YA=0}
| | | | | JENIS KELAMIN = P: TIDAK {TIDAK=9, YA=0}
| | | SUARA SAH CALEG ≤ 1468: YA {TIDAK=0, YA=3}
| SUARA SAH CALEG ≤ 1436.500: TIDAK {TIDAK=313, YA=0}
```

Nama : Wahyu Adi Nugroho

NIM : A11.2019.12310

Kelp. : 46UG

Dari deskripsi di atas, dapat diketahui bahwa :

- a. Atribut atau faktor yang paling berpengaruh adalah : **SUARA SAH CALEG, NO URUT CALEG.**
- b. Atribut atau faktor yang tidak berpengaruh adalah : **DAERAH PEMILIHAN, SUARA SAH PARTAI, JENIS KELAMIN.**

- **Penerapan / Deployment**

- a. Mbak PM akan melakukan strategi kampanye yang bersih dengan menawarkan program kerja yang pro rakyat sehingga diharapkan dia dapat mendapatkan suara yang mayoritas karena faktor yang menentukan keterpilihan dirinya sebagai anggota legislatif ialah surat suara caleg yang sah.
- b. Mbak PM juga akan mengusahakan untuk mendapatkan no urut caleg yang tingkat keterpilihannya tinggi kendati pun pada kenyataannya akan susah untuk mendapatkan no urut sesuai dengan keinginannya karena biasanya pemilihan no urut berdasarkan undian sehingga tergantung keberuntungan.
- c. Jika kemungkinan pemilihnya kurang dari 1436.5 (dibulatkan menjadi 1437) surat suara sah, maka sebaiknya Mbak PM tidak usah mencalonkan diri sebagai caleg karena pasti **TIDAK TERPILIH.**
- d. Untuk daerah pemilihan, mungkin Mbak PM perlu untuk membuat pemetaan daerah mana saja yang sekiranya kuat bagi dirinya terpilih.