

Laporan Percobaan Decision Tree (ID3)

Priscilla Vanny Amelia – 0206021810012

Dataset yang digunakan:

	anaemia	high_blood_pressure	sex	smoking	diabetes	\
count	299.000000	299.000000	299.000000	299.000000	299.000000	
mean	0.431438	0.351171	0.648829	0.32107	0.418060	
std	0.496107	0.478136	0.478136	0.46767	0.494067	
min	0.000000	0.000000	0.000000	0.00000	0.000000	
25%	0.000000	0.000000	0.000000	0.00000	0.000000	
50%	0.000000	0.000000	1.000000	0.00000	0.000000	
75%	1.000000	1.000000	1.000000	1.00000	1.000000	
max	1.000000	1.000000	1.000000	1.00000	1.000000	

	DEATH_EVENT
count	299.00000
mean	0.32107
std	0.46767
min	0.00000
25%	0.00000
50%	0.00000
75%	1.00000
max	1.00000

	anaemia	high_blood_pressure	sex	smoking	diabetes	DEATH_EVENT
0	0	1	1	0	0	heart_failure
1	0	0	1	0	0	heart_failure
2	0	0	1	1	0	heart_failure
3	1	0	1	0	0	heart_failure
4	1	0	0	0	1	heart_failure
...
294	0	1	1	1	1	none
295	0	0	0	0	0	none
296	0	0	0	0	1	none
297	0	0	1	1	0	none
298	0	0	1	1	0	none

299 rows × 6 columns

Hasil olah decision tree:

```
{ 'high_blood_pressure': {0: { 'smoking': {0: { 'diabetes': {0: { 'sex': { 'Female': { 'anaemia': {0: { 'anaemia': {0: 'none'}}},
1: { 'anaemia': {1: 'none'}}}},
'Male': { 'anaemia': {0: { 'anaemia': {0: 'none'}}},
1: { 'anaemia': {1: 'heart_failur
e'}}}}}},
1: { 'anaemia': {0: { 'sex': { 'Female': { 'anaemia': {0: 'none'}}},
'Male': { 'anaemia': {0: 'none'}}}},
1: { 'sex': { 'Female': { 'anaemia': {1: 'none'}}},
'Male': { 'anaemia': {1: 'none'}}}}}},
1: { 'sex': { 'Female': 'heart_failure',
'Male': { 'diabetes': {0: { 'anaemia': {0: { 'anaemia': {0: 'none'}}},
1: { 'anaemia': {1: 'none'}}}},
1: { 'anaemia': {0: { 'anaemia': {0: 'none'}}},
1: { 'anaemia': {1: 'none'}}}}}}}},
1: { 'smoking': {0: { 'diabetes': {0: { 'anaemia': {0: { 'sex': { 'Female': { 'anaemia': {0: 'none'}}},
'Male': { 'anaemia': {0: 'none'}}}},
1: { 'sex': { 'Female': { 'anaemia': {1: 'none'}}},
'Male': { 'anaemia': {1: 'none'}}}}}},
1: { 'sex': { 'Female': { 'anaemia': {0: { 'anaemia': {0: 'none'}}},
1: { 'anaemia': {1: 'heart_failur
e'}}}},
'Male': { 'anaemia': {0: { 'anaemia': {0: 'none'}}},
1: 'none'}}}}}},
1: { 'sex': { 'Female': { 'anaemia': {0: { 'diabetes': {0: 'none',
1: 'heart_failure'}}},
1: 'heart_failure'}}},
'Male': { 'diabetes': {0: { 'anaemia': {0: { 'anaemia': {0: 'none'}}},
1: { 'anaemia': {1: 'heart_failur
e'}}}},
1: { 'anaemia': {0: { 'anaemia': {0: 'heart_failur
e'}}}},
1: 'none'}}}}}}}}}
```

Kesimpulan:

Hasil pengolahan decision tree:

Orang yang cenderung terkena serangan jantung:

- Pria dengan anemia cenderung terkena serangan jantung (meskipun tidak tekanan darah tinggi, tidak merokok, dan tidak diabetes)
- Wanita dengan tekanan darah tinggi, perokok, diabetes, dan anemia, cenderung terkena serangan jantung
- Wanita dengan tekanan darah tinggi, perokok, diabetes cenderung terkena serangan jantung
- Wanita dengan tekanan darah tinggi, perokok, dan anemia cenderung terkena serangan jantung
- Pria dengan tekanan darah tinggi, perokok, dan anemia cenderung terkena serangan jantung
- Pria dengan tekanan darah tinggi, perokok, dan diabetes cenderung terkena serangan jantung

Terdapat anomaly saat menjalankan algoritma berupa inkonsistensi pada atribut anemia. Untuk mengatasinya, rekursif akan dibatasi hingga 6 kali perulangan, jika lebih dari 6 kali, maka keputusan akan diambil berdasarkan mayoritas.

Fitur yang paling berpengaruh terhadap variable target (Death event – karena serangan jantung) adalah Tekanan darah tinggi dan perokok/bukan, diikuti oleh Anemia dan Diabetes.