

LEMBAR TUGAS MAHASISWA

Disusun Untuk memenuhi UTS Data Mining

Dosen Pengajar Sri Erina Damayanti, S.T., M.Kom.



Nama	: Shaskia Putri Devi
NPM	: 21552011082
Kelas / Semester	: 221PA Semester 6
Program Studi Keahlian	: Teknik Informatika

PROGRAM STUDI TEKNIK INFORMATIKA

UNIVERSITAS TEKNOLOGI BANDUNG

TAHUN 2023 – 2024

Klasifikasi Tingkat Pendidikan Penduduk Desa di Jawa Barat Menggunakan Algoritma Naive Bayes dengan Aplikasi Orange

A. Latar Belakang

Dataset ini diperoleh dari situs web resmi Pemerintah Provinsi Jawa Barat, tepatnya dari <https://opendata.jabarprov.go.id/>. Dataset ini disusun oleh Dinas Pemberdayaan Masyarakat dan Desa dan diterbitkan secara tahunan. Dataset ini berisi informasi tentang tingkat pendidikan sebagian besar penduduk desa di Jawa Barat selama periode waktu tertentu.

Alasan mengapa klasifikasi dataset ini penting adalah untuk memahami pola dan karakteristik tingkat pendidikan penduduk desa di Jawa Barat. Dengan menggunakan algoritma klasifikasi seperti Naive Bayes, kita dapat memprediksi atau mengklasifikasikan tingkat pendidikan berdasarkan variabel-variabel tertentu yang tersedia dalam dataset.

Analisis ini dapat menjadi dasar untuk perencanaan kebijakan pendidikan di tingkat regional atau lokal. Selain itu, analisis ini berguna untuk mengevaluasi efektivitas program-program pendidikan yang sudah ada dan dapat memberikan bukti empiris yang kuat untuk mendukung pengambilan keputusan terkait dengan kebijakan pendidikan. Dengan menggunakan data dan analisis yang solid, pengambilan keputusan dapat lebih objektif dan efektif..

Jadi, melalui klasifikasi dataset ini dengan algoritma Naive Bayes menggunakan aplikasi Orange, kita dapat mendapatkan wawasan yang berharga tentang profil pendidikan penduduk desa di Jawa Barat dan mengambil langkah-langkah yang tepat untuk meningkatkan akses dan kualitas pendidikan di wilayah tersebut.

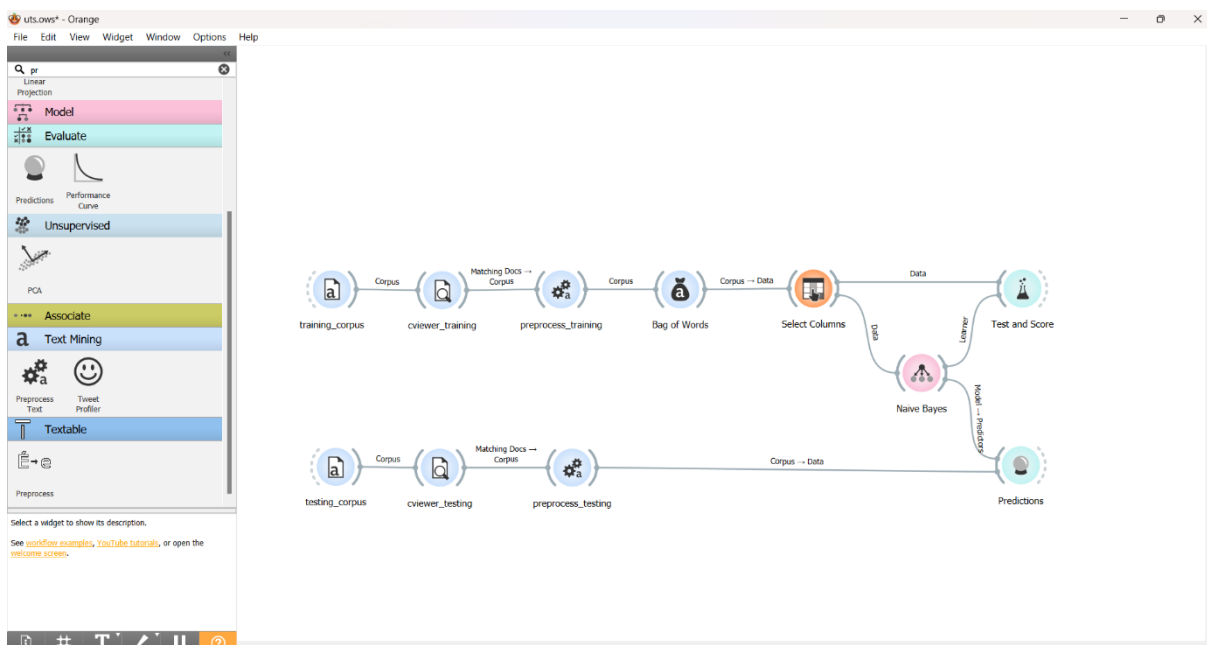
B. Proses data mining dengan menggunakan teknik data mining Klasifikasi

Variabel yang terdapat dalam dataset ini mencakup nama kecamatan, nama desa atau kelurahan, serta tingkat pendidikan.

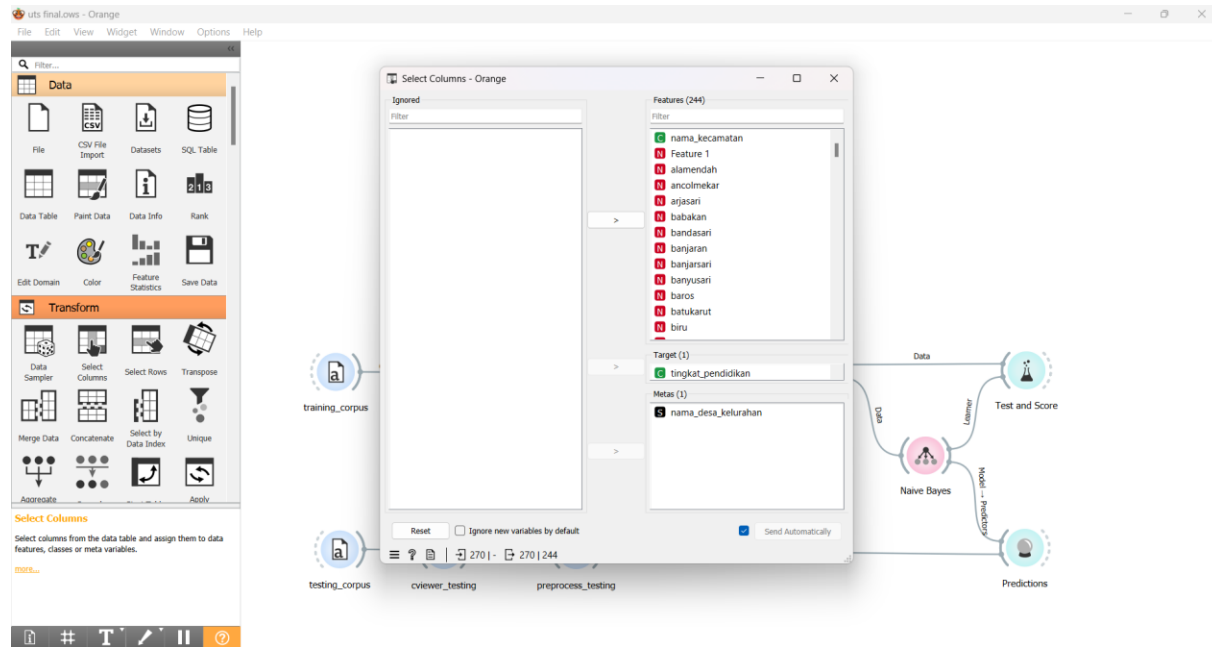
- Dataset

	A	B	C	D
1	nama_kecamatan	nama_desa_kelurahan	tingkat_pendidikan	
2	CILEUNYI	CILEUNYI KULON	TAMAT SMP/SEDERAJAT	
3	CILEUNYI	CILEUNYI WETAN	TAMAT SMA/SEDERAJAT	
4	CILEUNYI	CIMEKAR	TAMAT SMA/SEDERAJAT	
5	CILEUNYI	CINUNUK	TAMAT SMA/SEDERAJAT	
6	CILEUNYI	CIBIRU HILIR	TAMAT SMA/SEDERAJAT	
7	CILEUNYI	CIBIRU WETAN	TAMAT SMA/SEDERAJAT	
8	CIMENYAN	CIMENYAN	TAMAT SMP/SEDERAJAT	
9	CIMENYAN	MANDALAMEKAR	TAMAT SMP/SEDERAJAT	
10	CIMENYAN	CIKADUT	TAMAT SMA/SEDERAJAT	
11	CIMENYAN	CIBURIAL	TAMAT SD/SEDERAJAT	
12	CIMENYAN	SINDANGLAYA	TAMAT SMA/SEDERAJAT	
13	CIMENYAN	MEKARSALUYU	TAMAT SMP/SEDERAJAT	
14	CIMENYAN	MEKARMANIK	TAMAT SMP/SEDERAJAT	
15	CILENGKRANG	JATIENDAH	TAMAT SMA/SEDERAJAT	
16	CILENGKRANG	CILENGKRANG	TAMAT SD/SEDERAJAT	
17	CILENGKRANG	CIPANJALU	TAMAT SD/SEDERAJAT	
18	CILENGKRANG	MELATIWANGI	TAMAT SMP/SEDERAJAT	
19	CILENGKRANG	CIPOREAT	TAMAT SD/SEDERAJAT	
20	CILENGKRANG	GIRIMEKAR	TAMAT SD/SEDERAJAT	
21	BOJONGSOANG	LENGKONG	TAMAT SMA/SEDERAJAT	
22	BOJONGSOANG	BOJONGSOANG	TAMAT SMA/SEDERAJAT	
23	BOJONGSOANG	BUAHBATU	TAMAT SMA/SEDERAJAT	
24	BOJONGSOANG	CIPAGALO	TAMAT SMA/SEDERAJAT	
25	BOJONGSOANG	BOJONGSARI	TAMAT SMA/SEDERAJAT	
26	BOJONGSOANG	TEGALLUAR	TAMAT SMP/SEDERAJAT	
27	MARGAHAYU	MARGAHAYU TENGAH	TAMAT SMA/SEDERAJAT	
28	MARGAHAYU	MARGAHAYU SELATAN	TAMAT SMA/SEDERAJAT	
29	MARGAHAYU	SUKAMENAK	TAMAT SMA/SEDERAJAT	
30	MARGAHAYU	SAYATI	TAMAT SMA/SEDERAJAT	
31	MARGAASIH	MARGAASIH	TIDAK SEKOLAH	
32	MARGAASIH	LAGADAR	TAMAT SD/SEDERAJAT	
33	MARGAASIH	NANJUNG	TAMAT SMA/SEDERAJAT	

- Proses klasifikasi



- Atribut yang dijadikan target



C. Hasil Pengolahan Teknik Data Mining

- Hasil dari Predictions

Predictions - Orange		
Show probabilities for Classes known to the model		
	Naive Bayes	
		nama_kecamatan nama_desa_kelurahan
1	0.58 : 0.07 : 0.14 : 0.09 : 0.12 → TAMAT SARJANA/D1/D3/SEDERAJAT	CILEUNYI
2	0.56 : 0.03 : 0.20 : 0.09 : 0.12 → TAMAT SARJANA/D1/D3/SEDERAJAT	CILEUNYI KULON
3	0.68 : 0.04 : 0.08 : 0.05 : 0.14 → TAMAT SARJANA/D1/D3/SEDERAJAT	CILEUNYI WETAN
4	0.68 : 0.04 : 0.08 : 0.05 : 0.14 → TAMAT SARJANA/D1/D3/SEDERAJAT	CIMEKAR
5	0.69 : 0.04 : 0.02 : 0.10 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	CINUNUK
6	0.71 : 0.04 : 0.04 : 0.05 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	MANDALAMEKAR
7	0.69 : 0.09 : 0.02 : 0.05 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	CIKADUT
8	0.71 : 0.04 : 0.04 : 0.05 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	CIBURIAL
9	0.69 : 0.04 : 0.02 : 0.10 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	SINDANGLAYA
10	0.70 : 0.09 : 0.04 : 0.03 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	MEKARSALUYU
11	0.19 : 0.33 : 0.15 : 0.16 : 0.16 → TAMAT SD/SEDERAJAT	JATIENDAH
12	0.65 : 0.17 : 0.02 : 0.02 : 0.14 → TAMAT SARJANA/D1/D3/SEDERAJAT	CILENGKRANG
13	0.69 : 0.09 : 0.02 : 0.05 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	CILENGKRANG
14	0.65 : 0.17 : 0.02 : 0.02 : 0.14 → TAMAT SARJANA/D1/D3/SEDERAJAT	CIPANJALU
15	0.70 : 0.04 : 0.08 : 0.03 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	MELATIWANGI
16	0.20 : 0.17 : 0.31 : 0.16 : 0.17 → TAMAT SMA/SEDERAJAT	GIRIMEKAR
17	0.70 : 0.04 : 0.08 : 0.03 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	LENGKONG
18	0.70 : 0.04 : 0.08 : 0.03 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	BOJONGSOANG
19	0.17 : 0.14 : 0.41 : 0.14 : 0.14 → TAMAT SMA/SEDERAJAT	BOJONGSOANG
20	0.67 : 0.04 : 0.12 : 0.03 : 0.14 → TAMAT SARJANA/D1/D3/SEDERAJAT	BUAHBATU
21	0.67 : 0.04 : 0.12 : 0.03 : 0.14 → TAMAT SARJANA/D1/D3/SEDERAJAT	CIPAGALO
22	0.16 : 0.14 : 0.13 : 0.27 : 0.30 → TIDAK SEKOLAH	MARGAHAYU
23	0.60 : 0.04 : 0.05 : 0.04 : 0.27 → TAMAT SARJANA/D1/D3/SEDERAJAT	MARGAHAYU TENGAH
24	0.61 : 0.04 : 0.03 : 0.05 : 0.27 → TAMAT SARJANA/D1/D3/SEDERAJAT	MARGAHAYU SELATAN
25	0.70 : 0.04 : 0.08 : 0.03 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	SUKAMENAK
26	0.20 : 0.17 : 0.31 : 0.16 : 0.17 → TAMAT SMA/SEDERAJAT	MARGAASIH
27	0.71 : 0.04 : 0.04 : 0.05 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	MARGAASIH
28	0.68 : 0.04 : 0.08 : 0.05 : 0.14 → TAMAT SARJANA/D1/D3/SEDERAJAT	NANIJUNG
29	0.20 : 0.17 : 0.31 : 0.16 : 0.17 → TAMAT SMA/SEDERAJAT	MEKARRAHAYU
30	0.89 : 0.00 : 0.06 : 0.00 : 0.05 → TAMAT SARJANA/D1/D3/SEDERAJAT	KATAPANG
31	0.90 : 0.01 : 0.04 : 0.00 : 0.05 → TAMAT SARJANA/D1/D3/SEDERAJAT	SANGKANHURIP
32	0.68 : 0.04 : 0.08 : 0.05 : 0.14 → TAMAT SARJANA/D1/D3/SEDERAJAT	KATAPANG
		KATAPANG
		GANDASARI
		SUKAMUKTI
		DAYEUKOLOT
		DAYEUKOLOT
		CANGKUANG WETAN
		CANGKUANG KULON
		KIKAPIRA

Predictions - Orange		
Show probabilities for Classes known to the model		
	Naive Bayes	
		nama_kecamatan nama_desa_kelurahan
84	0.19 : 0.33 : 0.15 : 0.16 : 0.16 → TAMAT SD/SEDERAJAT	SOLOKANERUK
85	0.20 : 0.17 : 0.31 : 0.16 : 0.17 → TAMAT SMA/SEDERAJAT	PANYADAP
86	0.19 : 0.17 : 0.15 : 0.32 : 0.16 → TAMAT SMP/SEDERAJAT	PASEH
87	0.19 : 0.33 : 0.15 : 0.16 : 0.16 → TAMAT SD/SEDERAJAT	CIGENTUR
88	0.19 : 0.33 : 0.15 : 0.16 : 0.16 → TAMAT SD/SEDERAJAT	PASEH
89	0.20 : 0.17 : 0.31 : 0.16 : 0.17 → TAMAT SMA/SEDERAJAT	CIPEDES
90	0.71 : 0.04 : 0.04 : 0.05 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	LOA
91	0.70 : 0.09 : 0.04 : 0.03 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	PASEH
92	0.71 : 0.04 : 0.04 : 0.05 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	CUAGRA
93	0.20 : 0.17 : 0.31 : 0.16 : 0.17 → TAMAT SMA/SEDERAJAT	IBUN
94	0.88 : 0.02 : 0.02 : 0.01 : 0.07 → TAMAT SARJANA/D1/D3/SEDERAJAT	IBUN
95	0.70 : 0.04 : 0.08 : 0.03 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	LAKSANA
96	0.88 : 0.02 : 0.02 : 0.01 : 0.07 → TAMAT SARJANA/D1/D3/SEDERAJAT	IBUN
97	0.20 : 0.17 : 0.31 : 0.16 : 0.17 → TAMAT SMA/SEDERAJAT	DUKUH
98	0.67 : 0.04 : 0.04 : 0.10 : 0.14 → TAMAT SARJANA/D1/D3/SEDERAJAT	TALUN
99	0.67 : 0.08 : 0.08 : 0.02 : 0.14 → TAMAT SARJANA/D1/D3/SEDERAJAT	SOREANG
100	0.71 : 0.04 : 0.04 : 0.05 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	SOREANG
101	0.70 : 0.09 : 0.04 : 0.03 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	SADU
102	0.62 : 0.04 : 0.04 : 0.02 : 0.28 → TAMAT SARJANA/D1/D3/SEDERAJAT	PANYIRAPAN
103	0.20 : 0.17 : 0.31 : 0.16 : 0.17 → TAMAT SMA/SEDERAJAT	SUKAJADI
104	0.70 : 0.04 : 0.08 : 0.03 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	PASIRJAMBU
105	0.62 : 0.04 : 0.04 : 0.02 : 0.28 → TAMAT SARJANA/D1/D3/SEDERAJAT	PASIRJAMBU
106	0.19 : 0.33 : 0.15 : 0.16 : 0.16 → TAMAT SD/SEDERAJAT	CIBODAS
107	0.19 : 0.33 : 0.15 : 0.16 : 0.16 → TAMAT SD/SEDERAJAT	PASIRJAMBU
108	0.19 : 0.17 : 0.15 : 0.32 : 0.16 → TAMAT SMP/SEDERAJAT	CIKONENG
109	0.62 : 0.04 : 0.18 : 0.02 : 0.13 → TAMAT SARJANA/D1/D3/SEDERAJAT	CUKANGGENTENG
110	0.64 : 0.08 : 0.09 : 0.05 : 0.14 → TAMAT SARJANA/D1/D3/SEDERAJAT	CISONDARI
111	0.62 : 0.04 : 0.18 : 0.02 : 0.13 → TAMAT SARJANA/D1/D3/SEDERAJAT	CIWIDEY
112	0.56 : 0.14 : 0.02 : 0.02 : 0.25 → TAMAT SARJANA/D1/D3/SEDERAJAT	LEBAKMUNCANG
113	0.68 : 0.09 : 0.04 : 0.05 : 0.14 → TAMAT SARJANA/D1/D3/SEDERAJAT	CIWIDEY
114	0.61 : 0.08 : 0.02 : 0.02 : 0.27 → TAMAT SARJANA/D1/D3/SEDERAJAT	NENGKELAN
115	0.69 : 0.09 : 0.02 : 0.05 : 0.15 → TAMAT SARJANA/D1/D3/SEDERAJAT	CIWIDEY
		PANUNDAAN
		PATENGAN
		SUKARESMI
		INDRAGIRI
		CILUNCAT
		NAGRAK
		BANDASARI
		JELEGONG
		JATISARI
		PAMEUNTASAN
		KOPO

- Hasil dari Test and Score

The screenshot displays the Orange Data Mining interface. On the left is a widget palette with categories like Linear, Projection, Model, Evaluate, Unsupervised, Associate, Text Mining, and Textable. The main workspace contains the 'Test and Score' widget. The widget's settings are configured with 'Cross validation' (3 folds, stratified), 'Random sampling' (10 repeats, 66% training set size), and 'Test on train data' selected. The 'Evaluation results for target' table shows the following metrics for the Naive Bayes model:

Model	AUC	CA	F1	Prec	Recall	MCC
Naive Bayes	0.925	0.070	0.095	0.578	0.070	0.184

Below the table, the 'Compare models by: Area under ROC curve' section shows a comparison between 'Naive Bayes' and 'Naive Ba...'. The bottom status bar indicates the widget size is 270x270.

D. Kesimpulan

Berdasarkan analisis, terlihat pola tingkat pendidikan penduduk di desa-desa Jawa Barat selama periode tertentu. Data menunjukkan variasi yang signifikan dalam tingkat pendidikan, dengan beberapa desa memiliki tingkat pendidikan yang lebih tinggi dari yang lain. Dengan memahami pola-pola tingkat pendidikan, kita dapat mengidentifikasi area-area yang memerlukan perhatian khusus, seperti peningkatan akses pendidikan atau kualitas pendidikan.

Hasil analisis ini dapat membantu pemerintah daerah merancang program-program pendidikan yang lebih efektif dan berkelanjutan, serta mengalokasikan sumber daya yang sesuai untuk memenuhi kebutuhan pendidikan di berbagai wilayah. Meskipun ada desa-desa dengan tingkat pendidikan yang tinggi, terdapat juga desa-desa dengan tingkat pendidikan yang rendah, menunjukkan adanya potensi untuk perbaikan dalam akses dan kualitas pendidikan di beberapa wilayah.

E. Referensi Pengambilan Data

<https://opendata.jabarprov.go.id/id/dataset/tingkat-pendidikan-sebagian-besar-penduduk-desa-berdasarkan-desakelurahan-di-jawa-barat>