

Analisis Kualitas Wine Menggunakan Support Vector Machine (SVM)

1. Pendahuluan

Aplikasi ini dibuat menggunakan Streamlit untuk menampilkan analisis kualitas wine berdasarkan dataset Wine Quality (WineQT.csv).

Algoritma yang digunakan adalah Support Vector Machine (SVM) dengan tiga kernel berbeda: Linear, Polynomial, dan Radial Basis Function (RBF).

Tujuan analisis adalah membandingkan performa masing-masing kernel berdasarkan Accuracy, Precision, Recall, dan F1-Score, serta menampilkan hasil analisis secara interaktif di aplikasi web Streamlit.

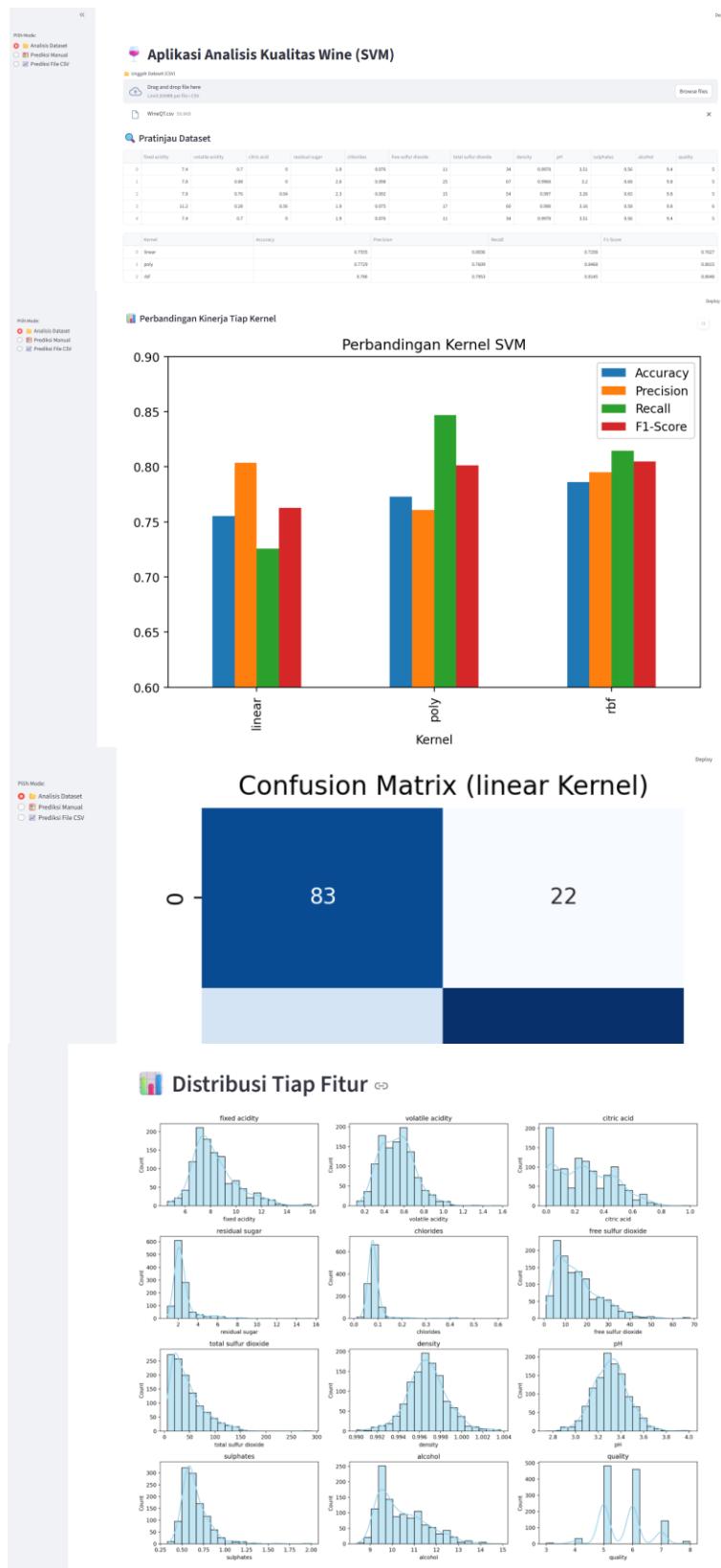
2. Deskripsi Program

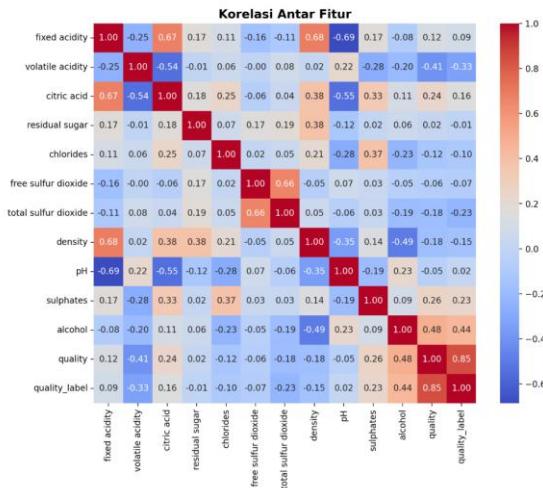
2.1 Struktur Program

Program terdiri dari dua kelas utama:

1. SVMWineModel → Mengelola preprocessing, pelatihan model, dan prediksi.
2. Visualizer → Menampilkan grafik distribusi, korelasi, perbandingan kernel, dan confusion matrix.

2.2 Tampilan Streamlit

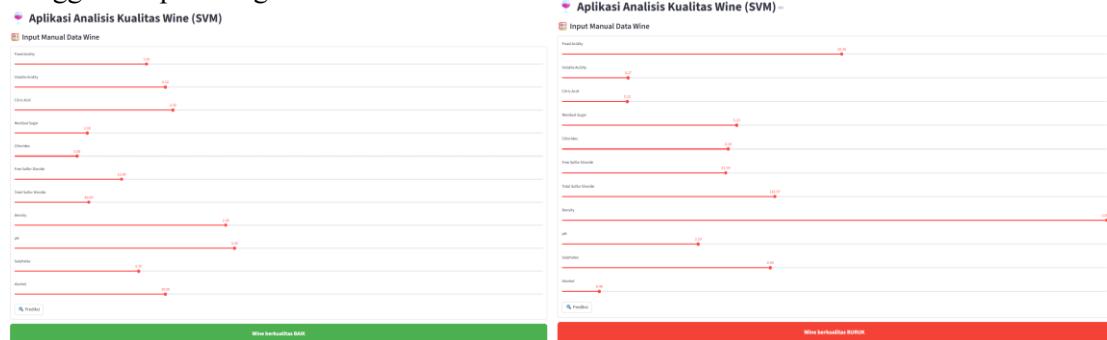




Setelah dataset diunggah, aplikasi akan menampilkan pratinjau dataset serta tabel hasil perbandingan metrik tiap kernel.

Menampilkan grafik perbandingan metrik kernel dan confusion matrix kernel terpilih.

Pengguna dapat mengisi fitur wine secara



manual melalui slider, kemudian aplikasi menampilkan hasil prediksi apakah wine termasuk baik atau buruk.

3. Penjelasan Alur Program

1. Load Dataset: Dataset WineQT.csv dimuat, kemudian kolom quality dikonversi menjadi dua kategori.
2. Preprocessing: Dataset dibagi menjadi data latih (80%) dan data uji (20%) menggunakan stratifikasi.
3. Training Model: Model SVM dilatih dengan tiga kernel: Linear, Polynomial, dan RBF.
4. Evaluasi Model: Menggunakan metrik Accuracy, Precision, Recall, dan F1-Score.
5. Visualisasi: Menampilkan grafik distribusi, korelasi, perbandingan metrik antar kernel, dan confusion matrix.

4. Hasil dan Pembahasan

Kernel	Accuracy	Precision	Recall	F1-Score
Linear	0.755459	0.803571	0.725806	0.762712
Polynomial	0.772926	0.760870	0.846774	0.801527
RBF	0.786026	0.795276	0.814516	0.804781

Secara keseluruhan, hasil pengujian menunjukkan bahwa **kernel RBF** memberikan performa terbaik dalam mengklasifikasikan kualitas wine dengan nilai *accuracy* 0.786 dan *F1-score* tertinggi sebesar 0.805. Hal ini menandakan bahwa RBF mampu menangkap pola non-linear pada data lebih baik dibanding kernel linear dan polynomial. Kernel **polynomial** menonjol dalam hal *recall* (0.847), menunjukkan sensitivitas tinggi terhadap data wine berkualitas baik, meskipun sedikit mengorbankan ketepatan. Sementara itu, **kernel linear** memiliki performa terendah di semua metrik karena tidak mampu memodelkan hubungan kompleks antar fitur. Dengan demikian, **RBF kernel direkomendasikan** sebagai pilihan utama karena keseimbangan antara akurasi, ketepatan, dan kemampuan generalisasi yang baik. **Analisis Tiap Kernel**

RBF (Radial Basis Function)

- Accuracy tertinggi (0.786) dan F1-Score terbaik (0.805).
- Performa seimbang antara *Precision* dan *Recall*, artinya model mampu mengenali wine baik tanpa banyak kesalahan positif atau negatif.
- Cocok untuk data non-linear karena RBF membuat batas keputusan melengkung, menyesuaikan pola kompleks pada dataset wine.

Kesimpulan: Kernel terbaik secara keseluruhan — memberikan keseimbangan antara akurasi dan generalisasi.

Polynomial

- Recall tertinggi (0.846) artinya model sangat sensitif terhadap wine baik sedikit yang terlewat.
- Namun precision lebih rendah (0.761) → model cenderung menganggap terlalu banyak data sebagai “baik”, sehingga lebih banyak false positive.
- Cocok jika tujuan utamamu adalah tidak melewatkannya wine berkualitas baik, meski dengan risiko kesalahan lebih banyak.

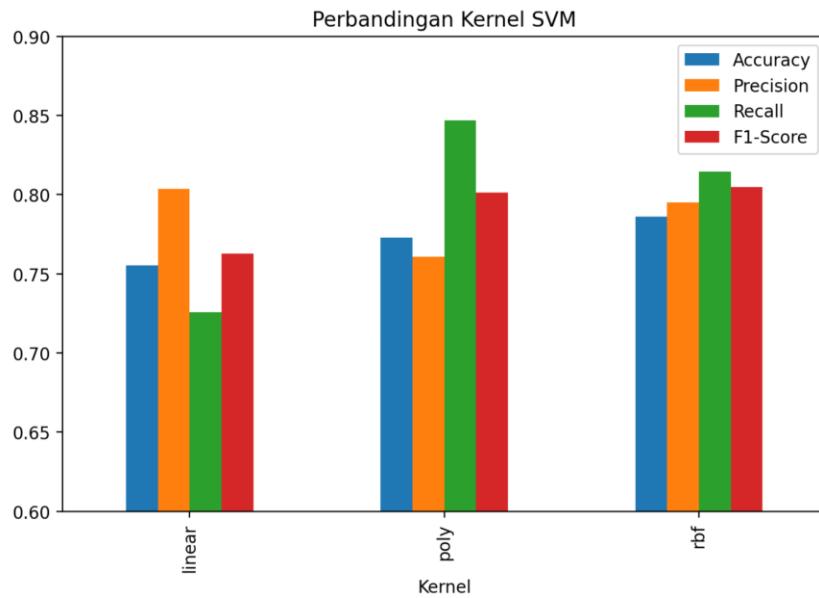
Kesimpulan: Baik untuk deteksi luas, tapi perlu kompromi pada ketepatan.

Linear

- Hasil paling rendah di semua metrik, terutama recall (0.726).
- Mengindikasikan hubungan antara fitur dan label tidak sepenuhnya linear.
- Kernel ini paling sederhana, jadi lebih cepat tapi kurang mampu menangkap kompleksitas data.

Kesimpulan: Cocok untuk baseline, tetapi tidak ideal untuk dataset dengan pola non-linear seperti wine.

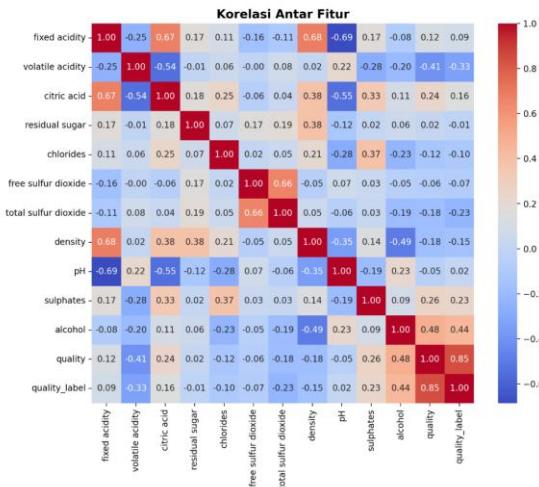
5. Visualisasi Hasil



Menampilkan diagram batang Accuracy, Precision, Recall, dan F1-Score tiap kernel.

	Kernel	Accuracy	Precision	Recall	F1-Score
0	linear	0.7555	0.8036	0.7258	0.7627
1	poly	0.7729	0.7609	0.8468	0.8015
2	rbf	0.786	0.7953	0.8145	0.8048

Kernel RBF menunjukkan klasifikasi paling seimbang antara label “baik” dan “buruk”.



Menunjukkan fitur-fitur yang paling mempengaruhi kualitas wine seperti alcohol, volatile acidity, dan density.