INSOLOGI: Jurnal Sains dan Teknologi
https://journal.literasisains.id/index.php/INSOLOGI
ISSN 2828-4984 (Media Online) | ISSN 2828-4992 (Media Cetak)
Vol. 4 No. 3 (Juni 2025) 656-668
DOI: 10.55123/insologi.v4i3.5436
Submitted: 28-04-2025 | Accepted: 16-05-2025 | Published: 15-06-2025

Analisis Kualitas Wine Menggunakan Machine Learning dengan Pendekatan SMOTE dan Seleksi Fitur

Triandes Sinaga[1], Kevin Bastian Sirait[2*], Jefri Junifer Pangaribuan[3], Okky Putra Barus[4], Romindo[5]
[1,2,3,4,5]Fakultas Teknologi Informasi, Sistem Informasi (Kampus Kota Medan),
Universitas Pelita Harapan, Jakarta, Indonesia
Email: [1]triandes.sinaga@uph.edu, [2*]kevin.sirait@uph.edu, [3]jefri.pangaribuan@uph.edu,
[4]okky.barus@uph.edu, [5]romindo@uph.edu

Abstract

Conventional wine quality assessment remains reliant on subjective expert judgment, which introduces potential bias and inconsistency in quality control processes. This study aims to develop an objective and automated machine learning-based classification model to enhance the accuracy of wine quality prediction. To address the issue of class imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was applied, along with ANOVA F-test-based feature selection to optimize model performance. The White Wine Quality dataset from the UCI Machine Learning Repository (4,898 samples, 11 numerical features) was utilized to evaluate five classification algorithms: Naïve Bayes, Decision Tree, Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). Before SMOTE application, the Random Forest model achieved an accuracy of only 67.55%. After implementing SMOTE and parameter tuning, the Random Forest (Tuned) model demonstrated the best performance with 90.29% accuracy, 89.99% precision, 90.29% recall, and 89,97%. % F1-score. Additionally, Decision Tree and KNN algorithms also exhibited notable improvements. SMOTE effectively balanced extreme minority class representations (quality levels 3 and 9). The most influential features in quality classification were alcohol content, density, and chlorides. These findings indicate that the proposed framework offers a reliable, objective, and scalable solution for automated wine quality control in industrial production environments.

Keywords: *Machine Learning, Wine Quality, SMOTE, Feature Selection, Random Forest, Classification.*

Abstrak

Penilaian kualitas wine secara konvensional masih bergantung pada penilaian subjektif para ahli, yang berpotensi menimbulkan bias dan inkonsistensi dalam proses kontrol kualitas. Penelitian ini bertujuan untuk mengembangkan model klasifikasi berbasis *machine learning* yang objektif dan otomatis guna meningkatkan akurasi prediksi kualitas wine. Untuk mengatasi permasalahan ketidakseimbangan kelas, diterapkan teknik *Synthetic Minority Over-sampling Technique* (SMOTE), disertai dengan seleksi fitur menggunakan anova f-test guna mengoptimalkan kinerja model. Dataset White Wine Quality dari UCI Machine Learning Repository (4.898 sampel, 11 fitur numerik) digunakan untuk mengevaluasi lima algoritma klasifikasi, yaitu: Naïve Bayes, Decision Tree, Random Forest, Support Vector Machine (SVM), dan K-Nearest Neighbors (KNN). Sebelum penerapan SMOTE, model Random Forest hanya mencapai akurasi sebesar 67,55%. Setelah penerapan SMOTE dan penyetelan parameter, model Random Forest (Tuned) menunjukkan performa terbaik dengan akurasi 90,29%, presisi 89,99%, recall 90,29%, dan skor F1 sebesar 89,97%. Selain itu, algoritma Decision Tree dan KNN juga menunjukkan peningkatan performa yang signifikan. SMOTE berhasil menyeimbangkan representasi kelas minoritas ekstrem (tingkat kualitas 3 dan 9). Analisis fitur menunjukkan bahwa kandungan alkohol, densitas, dan klorida merupakan prediktor paling berpengaruh dalam klasifikasi kualitas wine. Temuan ini menunjukkan bahwa kerangka kerja yang diusulkan menawarkan solusi yang andal, objektif, dan skalabel untuk kontrol kualitas wine secara otomatis di lingkungan produksi industri.

Kata Kunci: Machine Learning, Kualitas Wine, SMOTE, Seleksi Fitur, Random Forest, Klasifikasi.

1. PENDAHULUAN

Wine merupakan salah satu minuman fermentasi yang populer secara global, dengan kualitas yang dipengaruhi oleh berbagai faktor seperti varietas anggur, kondisi lingkungan, teknik produksi, dan metode penyimpanan. Kualitas wine tidak hanya menentukan preferensi konsumen tetapi juga berperan penting dalam daya saing industri wine secara keseluruhan (Conti et al., 2024; Croce et al., 2020). Khususnya, konsumen muda menunjukkan kepedulian yang meningkat terhadap kualitas dan pengalaman sensori yang ditawarkan oleh wine, menjadikan evaluasi kualitas produk sebagai aspek krusial dalam industri ini (Aich Satyabrata et al., 2019; Gupta, 2018).

Secara tradisional, penilaian kualitas wine dilakukan melalui evaluasi sensorik oleh ahli wine atau sommelier, yang menilai aspek-aspek seperti warna, aroma, rasa, dan tekstur menggunakan indera manusia (Cortez et al., 2009). Meskipun metode ini telah lama digunakan dan memberikan hasil yang akurat, penilaiannya bersifat subjektif dan memerlukan keterampilan khusus, sehingga menjadi tantangan bagi industri yang menginginkan konsistensi dalam kualitas produk (Gupta, 2018).

Sebagai alternatif, pendekatan berbasis teknologi seperti data mining dan machine learning semakin mendapatkan perhatian dalam evaluasi kualitas wine. Teknik-teknik ini memungkinkan identifikasi pola dalam data fisikokimia wine untuk membantu dalam klasifikasi kualitas secara lebih objektif dan efisien (Cortez et al., 2009; Zahedi et al., 2021). Berbagai algoritma machine learning telah diterapkan dalam prediksi kualitas wine, termasuk Naïve Bayes, Random Forest, Support Vector Machines (SVM), Decision Tree, dan K-Nearest Neighbors (K-NN)(Dahal et al., 2021; Kumar et al., 2020; Kurniasari et al., 2024; Kurniawan & Widi Nugroho, 2024).

Namun, penerapan machine learning dalam klasifikasi kualitas wine menghadapi tantangan terkait ketidakseimbangan kelas dalam dataset. Ketidakseimbangan ini dapat menyebabkan model cenderung memprediksi kelas mayoritas dan mengabaikan kelas minoritas, yang berpotensi mengurangi akurasi prediksi secara. Untuk mengatasi masalah ini, teknik Synthetic Minority Over-sampling Technique (SMOTE) dapat digunakan untuk menyeimbangkan distribusi kelas dengan menghasilkan contoh sintetis dari kelas minoritas (Chawla et al., 2002). Penerapan SMOTE dalam konteks klasifikasi kualitas wine telah menunjukkan peningkatan performa model dalam menghadapi data yang tidak seimbang (Dahal et al., 2021).

Namun, penerapan machine learning dalam klasifikasi kualitas wine menghadapi tantangan terkait ketidakseimbangan kelas dalam dataset. Ketidakseimbangan ini dapat menyebabkan model cenderung memprediksi kelas mayoritas dan mengabaikan kelas minoritas, yang berpotensi mengurangi akurasi prediksi. Untuk mengatasi masalah ini, teknik *Synthetic Minority Over-sampling Technique* (SMOTE) dapat digunakan untuk menyeimbangkan distribusi kelas dengan menghasilkan contoh sintetis dari kelas minoritas (Chawla et al., 2002). Penerapan SMOTE dalam konteks klasifikasi kualitas wine telah menunjukkan peningkatan performa model dalam menghadapi data yang tidak seimbang (Dahal et al., 2021). Bahkan, penelitian di bidang lain seperti diagnosis penyakit hati juga menunjukkan bahwa kombinasi SMOTE dan teknik reduksi dimensi seperti PCA dapat meningkatkan kualitas klasifikasi, meskipun terkadang berdampak pada perubahan akurasi model (Barus et al., 2022).

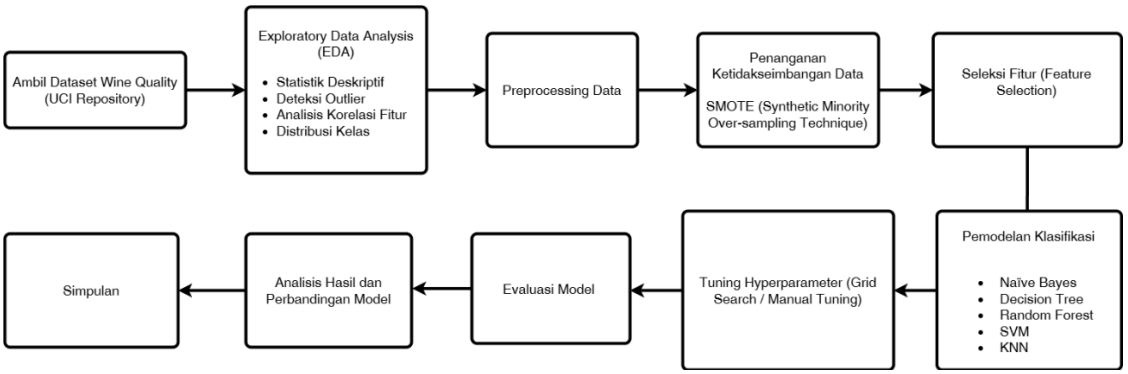
Selain itu, seleksi fitur juga menjadi aspek penting dalam meningkatkan kinerja model machine learning. Dengan mengidentifikasi dan memilih fitur-fitur yang paling relevan, model dapat mengurangi kompleksitas, meningkatkan interpretabilitas, dan menghindari overfitting (Gupta, 2018; Khakim et al., 2023). Berbagai teknik seleksi fitur, seperti correlation matrix, variance inflation factor (VIF), dan seleksi fitur berbasis

algoritma, telah diterapkan dalam penelitian klasifikasi kualitas wine untuk meningkatkan akurasi prediksi (Akinwande et al., 2015; Romindo et al., 2022; Vu et al., 2015)

Walaupun sejumlah penelitian sebelumnya telah membuktikan bahwa penerapan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) dan seleksi fitur secara terpisah dapat meningkatkan performa klasifikasi kualitas wine, penggabungan kedua pendekatan ini dalam satu model machine learning masih jarang dilakukan (Dahal et al., 2021; Kumar et al., 2020; Mor et al., 2022; Wayan et al., 2023). Padahal, pendekatan serupa di bidang lain menunjukkan hasil yang menjanjikan. Misalnya, (Sinaga et al., 2024) menunjukkan bahwa integrasi teknik lanjutan dalam pemrosesan data dan seleksi fitur secara signifikan meningkatkan akurasi dan stabilitas model. Temuan ini mengindikasikan bahwa strategi integratif SMOTE dan seleksi fitur juga berpotensi memberikan hasil yang lebih optimal dalam konteks klasifikasi kualitas wine. Oleh karena itu, penelitian ini bertujuan untuk mengisi kesenjangan tersebut dengan mengembangkan model prediksi kualitas wine melalui kombinasi SMOTE untuk penyeimbangan kelas dan seleksi fitur untuk mengidentifikasi atribut paling relevan. Diharapkan pendekatan ini mampu menghasilkan model yang tidak hanya akurat, tetapi juga efisien dan andal.

2. METODOLOGI PENELITIAN

Penelitian ini bertujuan untuk melakukan evaluasi komparatif beberapa algoritma klasifikasi dalam memprediksi kualitas wine dengan penerapan teknik feature selection dan penanganan ketidakseimbangan data. Metodologi penelitian terdiri dari beberapa tahapan utama yang digambarkan pada Gambar 1.



Gambar 1. Diagram alur

2.1 Pengambilan dan Eksplorasi Dataset

Dataset yang digunakan adalah Wine Quality Dataset (white wine) dari UCI Machine Learning Repository, terdiri dari 4898 sampel dengan 11 fitur numerik: fixed acidity, volatile acidity, citric acid, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, pH, sulphates, dan alcohol, serta satu atribut target berupa skor kualitas wine (rentang 3 - 9) (Cortez et al., 2009). Eksplorasi data awal dilakukan melalui analisis statistik deskriptif, visualisasi distribusi dan pencilan menggunakan boxplot, analisis distribusi kelas target, serta penghitungan korelasi antar fitur dengan koefisien Pearson untuk memahami hubungan fitur dan karakteristik data.

2.2 Pra-pemrosesan Data (Preprocessing)

Langkah pra-pemrosesan dimulai dengan pembersihan data untuk menangani nilai yang hilang (missing values) atau pencilan (outliers) yang signifikan, sesuai dengan rekomendasi (Olteanu et al., 2023) Selanjutnya, untuk memastikan model dapat

beroperasi secara optimal, semua fitur numerik distandarisasi menggunakan metode StandardScaler dari scikit-learn (de Amorim et al., 2022; Pedregosa et al., 2012).Proses ini penting karena banyak algoritma machine learning mengasumsikan bahwa data memiliki distribusi normal dengan rata-rata nol dan variansi satu, sehingga standarisasi membantu meningkatkan kinerja model secara keseluruhan.

2.3 Penanganan Ketidakseimbangan Kelas

Distribusi kelas target menunjukkan ketidakseimbangan dengan dominasi kelas skor 5 dan 6, sedangkan kelas ekstrim seperti 3 dan 9 sangat sedikit sampelnya. Untuk mengatasi hal ini, diterapkan teknik Synthetic Minority Over-sampling Technique (SMOTE) untuk mensintesis sampel baru pada kelas minoritas, sehingga distribusi kelas menjadi seimbang dan model tidak bias ke kelas mayoritas.

2.4 Seleksi Fitur (*Feature Selection*)

Seleksi fitur dilakukan dengan pendekatan univariate feature selection menggunakan uji statistik anova F-test (f_classif) untuk memilih fitur yang memiliki hubungan signifikan dengan target. Fitur dengan nilai p-value di bawah threshold tertentu (misalnya 0.05) dipertahankan agar model menjadi lebih sederhana, mengurangi risiko overfitting, dan mempercepat proses pelatihan.

2.5 Pemodelan Klasifikasi dan Penyetingan Hiperparameter

Lima algoritma klasifikasi diuji, yaitu Naïve Bayes, Decision Tree, Random Forest, Support Vector Machine (SVM), dan K-Nearest Neighbors (KNN). Semua model diimplementasikan menggunakan pustaka scikit-learn pada bahasa Python. Penyetelan hiperparameter dilakukan dengan kombinasi grid search (GridSearchCV) dan manual tuning untuk menemukan parameter terbaik yang memaksimalkan performa model.

2.6 Evaluasi Model

Model dievaluasi menggunakan metode stratified k-fold cross-validation dengan nilai *k* sebesar 10, sebagaimana direkomendasikan oleh (Raschka, 2018), guna mengurangi bias dan meningkatkan kestabilan hasil evaluasi. Evaluasi performa model dilakukan dengan menggunakan beberapa metrik, yaitu akurasi, presisi, recall, F1-score, dan Area Under the Curve (AUC). Penggunaan metrik-metrik ini dimaksudkan untuk memberikan penilaian yang komprehensif terhadap kinerja model, terutama dalam menghadapi masalah ketidakseimbangan kelas.

2.7 Analisis dan Interpretasi Hasil

Hasil evaluasi tiap algoritma dibandingkan secara komparatif untuk menentukan model terbaik dalam memprediksi kualitas wine. Analisis juga menilai dampak tiap tahap pipeline, pra-pemrosesan, penanganan ketidakseimbangan, dan seleksi fitur terhadap performa akhir model. Dengan demikian, penelitian ini memberikan rekomendasi tidak hanya dari sisi algoritma, tetapi juga strategi pemrosesan data yang efektif untuk sistem prediksi kualitas wine.

3. HASIL DAN PEMBAHASAN

3.1 Eksplorasi Data (Exploratory Data Analysis)

3.1.1 Informasi Umum Dataset

Dataset Wine Quality (white wine) yang diperoleh dari UCI Machine Learning Repository terdiri dari 4.898 sampel dengan 11 fitur numerik yang merepresentasikan karakteristik fisiko-kimia wine (Cortez et al., 2009). Fitur-fitur tersebut meliputi fixed

acidity, volatile acidity, citric acid, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, pH, sulphates, dan alcohol. Variabel target berupa quality merupakan skor kualitas wine dalam rentang 3 hingga 9. Struktur data awal dari dataset disajikan pada Tabel 1 untuk memberikan gambaran representatif mengenai distribusi nilai pada setiap atribut.

Tabel 1. Dataset White Wine

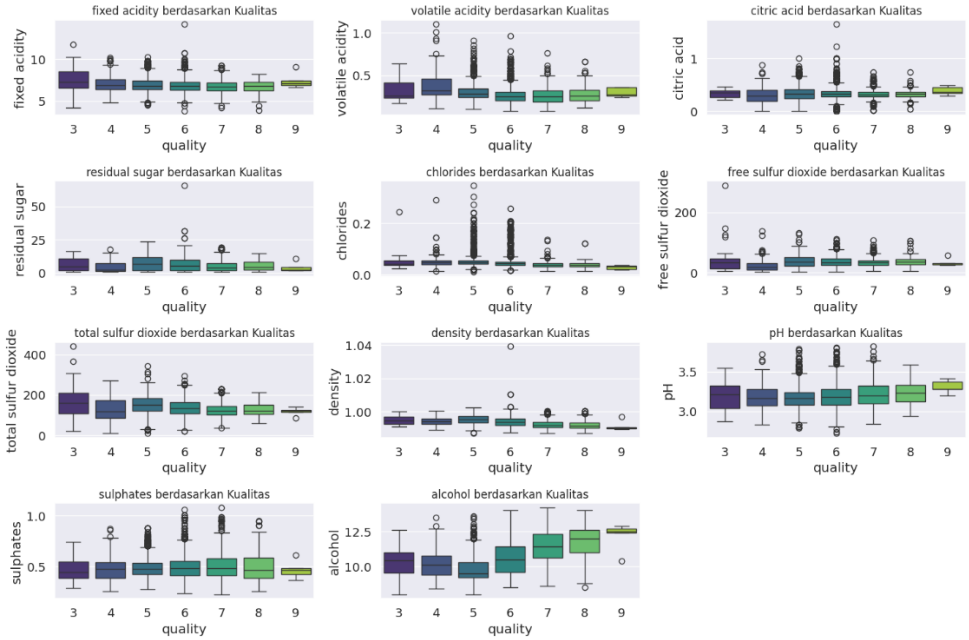
Fixed Acidity	Volatile Acidity	Citric Acid	Residual Sugar	Chlorides	Free SO_2	Total SO_2	Density	pH	Sulphates	Alcohol	Quality
7.0	0.27	0.36	20.7	0.045	45.0	170.0	1.0010	3.00	0.45	8.8	6
6.3	0.30	0.34	1.6	0.049	14.0	132.0	0.9940	3.30	0.49	9.5	6
8.1	0.28	0.40	6.9	0.050	30.0	97.0	0.9951	3.26	0.44	10.1	6
7.2	0.23	0.32	8.5	0.058	47.0	186.0	0.9956	3.19	0.40	9.9	6

3.1.2 Analisis Statistik Deskriptif

Analisis statistik deskriptif menunjukkan karakteristik distribusi setiap fitur dalam dataset. Fitur *fixed acidity* memiliki nilai rata-rata 6.86 dengan standar deviasi 0.844, menunjukkan distribusi yang relatif normal. Fitur *residual sugar* menampilkan variabilitas tertinggi dengan nilai maksimum mencapai 65.8, yang mengindikasikan adanya *outlier* signifikan. Variabel target *quality* memiliki rata-rata 5.88 dengan distribusi terpusat pada nilai 5 dan 6. Statistik lengkap untuk seluruh fitur disajikan pada Tabel 2, sedangkan deteksi *outlier* melalui visualisasi boxplot ditampilkan pada Gambar 2.

Tabel 2. Statistik Deskriptif Dataset *Wine Quality*

Fitur	Count	Mean	Std Dev	Min	25%	50%	75%	Max
Fixed Acidity	4898	6.855	0.844	3.80	6.30	6.80	7.30	14.20
Volatile Acidity	4898	0.278	0.101	0.08	0.21	0.26	0.32	1.10
Citric Acid	4898	0.334	0.121	0.00	0.27	0.32	0.39	1.66
Residual Sugar	4898	6.391	5.072	0.60	1.70	5.20	9.90	65.80
Chlorides	4898	0.046	0.022	0.009	0.036	0.043	0.050	0.346
Free SO_2	4898	35.31	17.01	2.00	23.00	34.00	46.00	289.00
Total SO_2	4898	138.36	42.50	9.00	108.0	134.0	167.0	440.00
Density	4898	0.994	0.003	0.987	0.992	0.994	0.996	1.039
pH	4898	3.188	0.151	2.72	3.09	3.18	3.28	3.82
Sulphates	4898	0.490	0.114	0.22	0.41	0.47	0.55	1.08
Alcohol	4898	10.514	1.231	8.00	9.50	10.40	11.40	14.20
Quality	4898	5.878	0.886	3.00	5.00	6.00	6.00	9.00



Gambar 2. Deteksi Outlier menggunakan Boxplot untuk Setiap Fitur

3.1.3 Analisis Distribusi Kelas dan Missing Values

Pemeriksaan terhadap missing values menunjukkan bahwa dataset tidak memiliki nilai yang hilang, sehingga dapat langsung diproses ke tahap selanjutnya. Analisis distribusi kelas target menunjukkan ketidakseimbangan yang signifikan, sebagaimana ditampilkan pada Tabel 3. Kelas quality 6 mendominasi dengan 44.88% dari total data, diikuti kelas 5 (29.75%) dan kelas 7 (17.96%). Sebaliknya, kelas ekstrim seperti 3 dan 9 hanya memiliki representasi minimal (0.41% dan 0.10%). Ketidakseimbangan ini menjadi perhatian utama dalam tahap preprocessing untuk mencegah bias model terhadap kelas mayoritas.

Tabel 3. Distribusi Kelas Target (Quality)

Kelas Quality	Jumlah Data	Persentase (%)
3	20	0.41
4	163	3.33
5	1,457	29.75
6	2,198	44.88
7	880	17.96
8	175	3.57
9	5	0.10

3.2 Pra-pemrosesan Data (Preprocessing)

3.2.1 Standardisasi Fitur

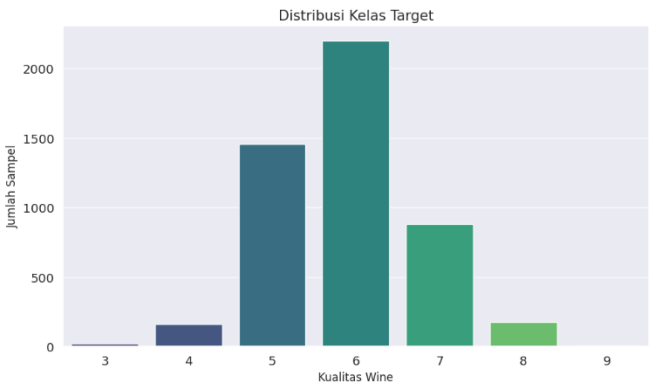
Mengingat tidak ditemukan missing values dalam dataset, tahap preprocessing difokuskan pada standardisasi fitur menggunakan Standard Scaler dari scikit-learn. Proses ini memastikan bahwa seluruh fitur memiliki skala yang seragam (mean = 0, std = 1), sehingga algoritma machine learning dapat beroperasi secara optimal tanpa bias terhadap fitur dengan skala nilai yang lebih besar.

3.2.2 Penanganan Outlier

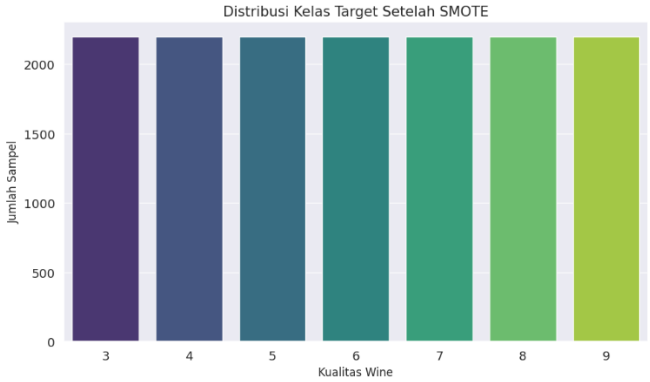
Berdasarkan analisis boxplot pada Gambar 2, ditemukan outlier signifikan pada beberapa fitur, terutama residual sugar, free SO_2, dan total SO_2. Namun, outlier ini dipertahankan karena masih dalam rentang nilai yang realistis untuk karakteristik wine dan tidak menunjukkan indikasi kesalahan pengukuran.

3.3 Penanganan Ketidakseimbangan Kelas

Ketidakseimbangan kelas yang teridentifikasi pada Tabel 3 ditangani menggunakan teknik Synthetic Minority Over-sampling Technique (SMOTE). Distribusi kelas sebelum penerapan SMOTE ditunjukkan pada Gambar 5, sedangkan hasil setelah penyeimbangan ditampilkan pada Gambar 6. Setelah penerapan SMOTE, setiap kelas memiliki jumlah sampel yang seimbang (2.198 sampel per kelas), dengan total dataset meningkat dari 4.898 menjadi 15.386 sampel.



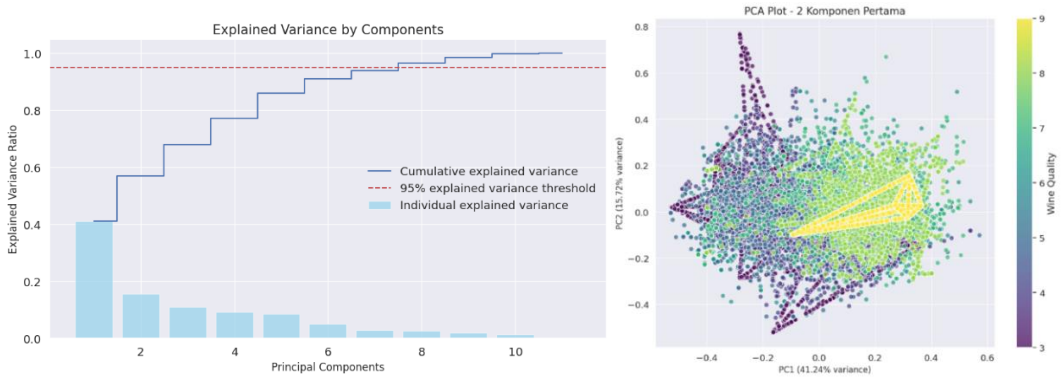
Gambar 5. Distribusi Kelas Target Sebelum SMOTE



Gambar 6. Distribusi Kelas Target Setelah Penerapan SMOTE.

3.4 Seleksi Fitur (Feature Selection)

Seleksi fitur dilakukan menggunakan pendekatan univariate feature selection dengan uji statistik ANOVA F-test (f_classif). Hasil analisis menunjukkan bahwa fitur alcohol, density, dan chlorides memiliki nilai F-score tertinggi dan p-value < 0.05, mengindikasikan hubungan signifikan dengan variabel target. Berdasarkan threshold p-value = 0.05, fitur-fitur dengan kontribusi statistik terbesar dipertahankan untuk tahap pemodelan. Sebagai tambahan, dilakukan analisis Principal Component Analysis (PCA) untuk reduksi dimensi, yang menunjukkan bahwa 8 komponen utama mampu menjelaskan 95% total varians data, sebagaimana ditampilkan pada Gambar 7.



Gambar 7. Analisis Komponen Utama (PCA) untuk Reduksi Dimensi

3.5 Pemodelan Klasifikasi dan Penyetelan Hiperparameter

Lima algoritma klasifikasi diimplementasikan dan dievaluasi: Naïve Bayes, Decision Tree, Random Forest, Support Vector Machine (SVM), dan K-Nearest Neighbors (KNN). Setiap algoritma menggunakan pustaka scikit-learn dengan penyetelan hiperparameter melalui GridSearchCV untuk mengoptimalkan performa. Hasil evaluasi awal pada data uji menunjukkan performa yang bervariasi, sebagaimana disajikan pada Tabel 4 merupakan tabel Hasil Performa Model sebelum dilakukan SMOTE dan Seleksi fitur. Dan pada Tabel 5 merupakan tabel evaluasi model pada data uji setelah dilakukan SMOTE dan Seleksi Fitur.

Tabel 4. Hasil Evaluasi Model sebelum SMOTE dan Fitur Seleksi.

No	Model	Akurasi	Presisi	Recall	Skor F1
1	Random Forest	0,675510	0,530549	0,387662	0,427719
2	Decision Tree	0,593878	0,359910	0,381524	0,369494
3	SVM	0,566327	0,399639	0,237097	0,240917
4	KNN	0,526531	0,341811	0,287819	0,306020
5	Naive Bayes	0,472449	0,306475	0,303183	0,285070

Tabel 5. Hasil Evaluasi Model setelah SMOTE dan Fitur Seleksi.

Algoritma	Accuracy	Precision	Recall	F1-Score
Naïve Bayes	0.4831	0.4858	0.4831	0.4626
Random Forest	0.9019	0.8992	0.9019	0.8997
SVM	0.6758	0.6628	0.6759	0.6653
Decision Tree	0.8298	0.8279	0.8298	0.8287
K-Nearest Neighbors	0.8203	0.8081	0.8204	0.8090

Sebelum penerapan metode SMOTE dan seleksi fitur, evaluasi lima model klasifikasi menunjukkan bahwa Random Forest memiliki performa terbaik dengan akurasi sebesar 67,55%, sedangkan Naïve Bayes menunjukkan kinerja terendah dengan akurasi 47,24%, yang kemungkinan disebabkan oleh asumsi independensi fitur yang tidak sesuai dengan karakteristik data wine. Setelah dilakukan penyeimbangan kelas menggunakan SMOTE dan pemilihan fitur berbasis anova F-test, kinerja seluruh model mengalami peningkatan signifikan. Random Forest tetap menunjukkan performa paling unggul dengan akurasi 90,19%, presisi 89,92%, recall 90,19%, dan F1-score 89,98%. Model Decision Tree dan K-Nearest Neighbors juga mencatatkan peningkatan performa yang substansial, sedangkan Naïve Bayes tetap menunjukkan hasil terendah dengan akurasi 48,31%. Hasil ini menegaskan bahwa penggunaan teknik penyeimbangan data dan seleksi fitur mampu meningkatkan efektivitas model klasifikasi dalam memprediksi kualitas wine, serta menjadikan pendekatan ini sebagai solusi yang andal dan objektif untuk sistem kontrol kualitas otomatis di industri produksi wine.

3.6 Evaluasi Model dengan Stratified K-Fold Cross-Validation

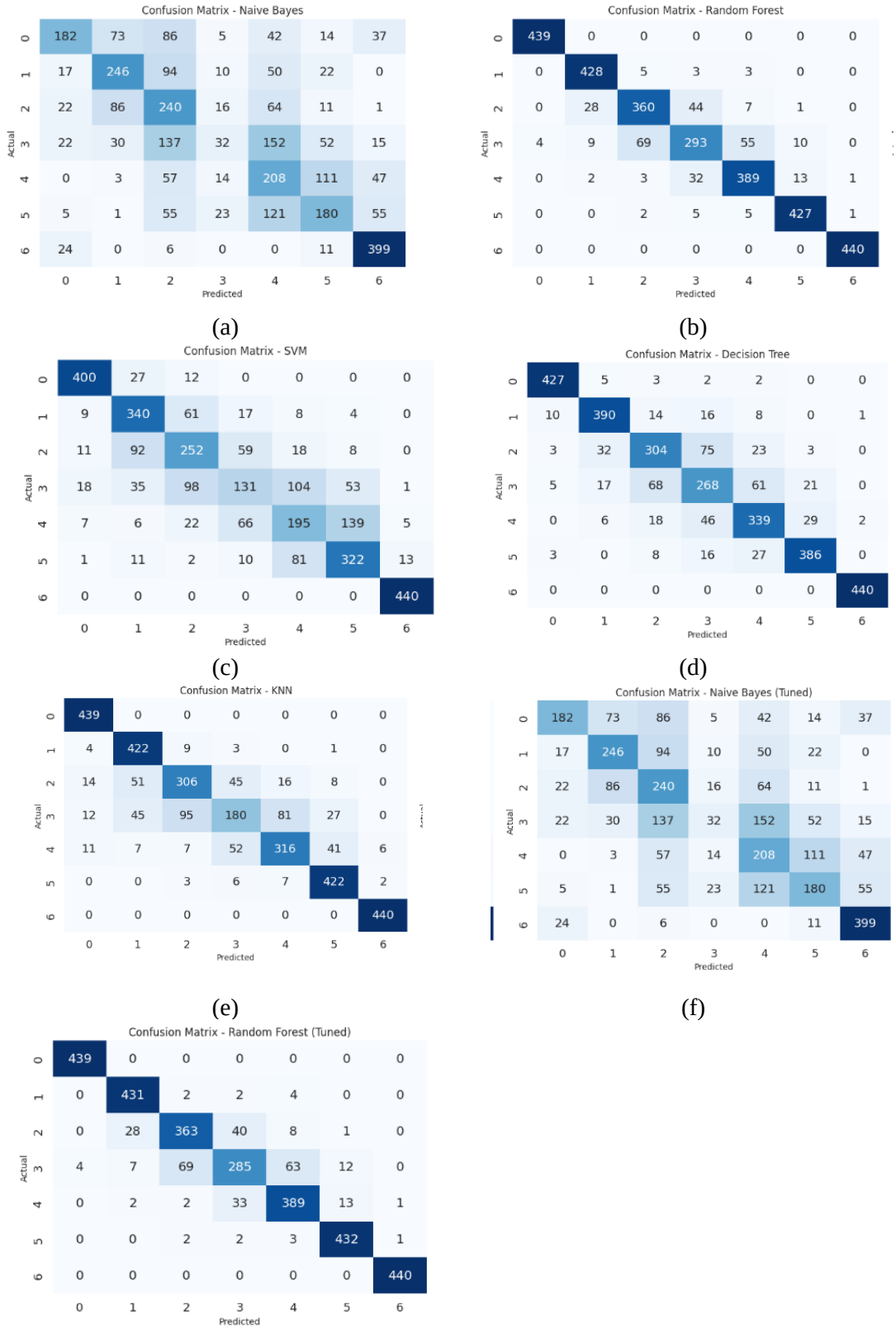
Untuk memastikan kestabilan dan generalisasi model, dilakukan evaluasi menggunakan stratified 10-fold cross-validation. Hasil evaluasi yang disajikan pada Tabel 6 mengonfirmasi konsistensi performa setiap algoritma. Random Forest mempertahankan akurasi tertinggi (89.17%) dengan standar deviasi rendah (0.0053), menunjukkan stabilitas model yang baik.

Tabel 6. Hasil Stratified 10-Fold Cross-Validation

Algoritma	Rata-rata Akurasi	Standar Deviasi
Naïve Bayes	0.4730	0.0066
Random Forest	0.8917	0.0053
SVM	0.6701	0.0049
Decision Tree	0.8252	0.0074
K-Nearest Neighbors	0.8123	0.0081

3.6.1 Analisis Confusion Matrix

Performa klasifikasi setiap algoritma dievaluasi melalui confusion matrix yang ditampilkan pada Gambar 8. Random Forest, baik versi standar maupun tuned, menunjukkan akurasi klasifikasi tertinggi pada seluruh kelas, terutama kelas mayoritas (5 dan 6). Naïve Bayes mengalami kesulitan signifikan dalam klasifikasi kelas 6, dengan tingkat misklasifikasi tinggi ke kelas 4 dan 5. Algoritma lainnya (KNN, SVM, Decision Tree) menunjukkan performa moderate dengan kesalahan klasifikasi yang lebih sering terjadi pada kelas minoritas.



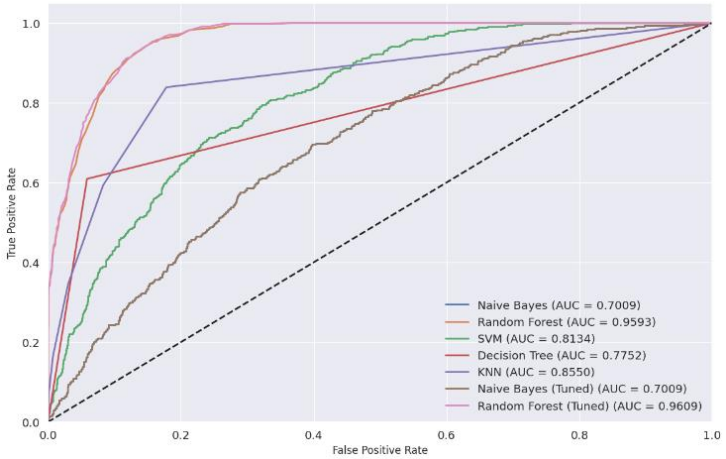
(a) (b) (c) (d) (e) (f) (g)

Gambar 8. Confusion Matrix untuk Setiap Algoritma Klasifikasi.

(a) Naïve Bayes, (b) Random Forest, (c) Support Vector Machine (SVM),(d) Decision Tree,(e) K-Nearest Neighbors (KNN), (f) Naïve Bayes (Tuned),(g) Random Forest (Tuned)

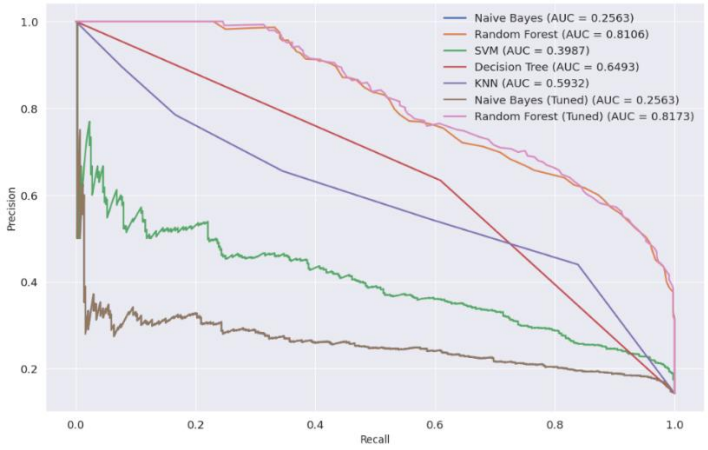
3.6.2 Analisis ROC Curve dan Precision-Recall Curve

Evaluasi komprehensif dilakukan menggunakan ROC Curve dan Precision-Recall Curve dengan fokus pada kelas dominan (kelas 6). Analisis ROC yang ditampilkan pada Gambar 9 menunjukkan bahwa Random Forest (Tuned) mencapai AUC tertinggi (0.9609), diikuti Random Forest standar (0.9593) dan KNN (0.8550). Naïve Bayes menunjukkan performa terendah dengan AUC 0.7009.



Gambar 9. ROC Curve untuk Evaluasi Performa Model

Precision-Recall Curve pada Gambar 10 memberikan evaluasi mendalam terhadap kemampuan model dalam mempertahankan precision seiring peningkatan recall, khususnya pada kondisi ketidakseimbangan kelas. Random Forest (Tuned) dan Random Forest standar menunjukkan AUC tertinggi (0.8173 dan 0.8106), sementara Naïve Bayes mencatat AUC terendah (0.2563). Hasil ini menegaskan superioritas ensemble learning dalam menangani kompleksitas data wine.



Gambar 10. Precision-Recall Curve untuk Analisis Ketidakseimbangan Kelas

3.7 Analisis dan Interpretasi Hasil

3.7.1 Perbandingan Performa Algoritma

Hasil evaluasi komprehensif menunjukkan bahwa Random Forest secara konsisten memberikan performa terbaik di seluruh metrik evaluasi. Keunggulan ini dapat diatribusikan pada kemampuan ensemble learning dalam mengatasi overfitting, menangani noise, dan mengakomodasi interaksi kompleks antar fitur. Sebaliknya, Naïve Bayes menunjukkan performa terendah, yang konsisten dengan keterbatasan asumsi independensi fitur pada data dengan korelasi antar variabel yang tinggi.

3.7.2 Dampak Tahapan Preprocessing

Penerapan SMOTE terbukti efektif dalam mengatasi ketidakseimbangan kelas, yang tercermin dari peningkatan performa klasifikasi pada kelas minoritas. Standardisasi fitur menggunakan StandardScaler memberikan kontribusi signifikan terhadap stabilitas konvergensi algoritma, terutama untuk SVM dan KNN yang sensitif terhadap skala fitur.

3.7.3 Efektivitas Seleksi Fitur

Univariate feature selection dengan ANOVA F-test berhasil mengidentifikasi fitur-fitur yang paling berkontribusi terhadap prediksi kualitas wine. Fitur alcohol, density, dan chlorides yang terpilih sejalan dengan temuan analisis korelasi Pearson, memberikan validasi terhadap konsistensi metodologi seleksi fitur.

3.7.4 Rekomendasi Model dan Strategi Preprocessing

Berdasarkan evaluasi komprehensif, Random Forest direkomendasikan sebagai algoritma optimal untuk prediksi kualitas wine dengan konfigurasi hyperparameter hasil tuning. Strategi preprocessing yang direkomendasikan meliputi: (1) penerapan SMOTE untuk mengatasi ketidakseimbangan kelas, (2) standardisasi fitur menggunakan StandardScaler, dan (3) seleksi fitur berbasis ANOVA F-test untuk efisiensi komputasi dan pencegahan overfitting. Temuan penelitian ini memberikan kontribusi praktis bagi industri wine dalam mengembangkan sistem prediksi kualitas yang akurat dan robust, serta memberikan wawasan metodologis bagi penelitian machine learning pada domain dengan karakteristik ketidakseimbangan kelas.

4. KESIMPULAN

Penelitian ini berhasil mengevaluasi kinerja berbagai algoritma klasifikasi dalam memprediksi kualitas wine dengan menerapkan pendekatan SMOTE untuk penanganan ketidakseimbangan kelas dan teknik seleksi fitur berbasis anova F-test. Hasil analisis menunjukkan bahwa ketidakseimbangan kelas yang signifikan pada dataset Wine Quality (white wine) dapat berdampak pada akurasi dan objektivitas model prediksi, namun dapat diatasi secara efektif dengan teknik SMOTE.

Penerapan seleksi fitur membantu menyederhanakan model, mengurangi risiko overfitting, serta meningkatkan efisiensi proses pelatihan dengan tetap mempertahankan fitur-fitur yang memiliki korelasi signifikan terhadap variabel target. Selain itu, pendekatan evaluasi menggunakan stratified k-fold cross-validation dan metrik performa yang beragam (akurasi, presisi, recall, F1-score, dan AUC) memberikan penilaian yang komprehensif terhadap kinerja masing-masing algoritma.

Secara umum, kombinasi strategi pra-pemrosesan, penanganan ketidakseimbangan data, dan seleksi fitur terbukti meningkatkan performa model klasifikasi. Hasil ini memberikan rekomendasi bahwa pendekatan berbasis SMOTE dan feature selection dapat menjadi strategi yang efektif dalam pengembangan sistem prediksi kualitas wine berbasis *machine learning.*

REFERENCES

Aich Satyabrata, Abdulhakim Al-Absi, A., Lee Hui, K., & Sain, M. (2019). *Prediction of Quality for Different Type of Wine based on Different Feature Sets Using Supervised Machine Learning Techniques*. IEEE.

Akinwande, M. O., Dikko, H. G., & Samson, A. (2015). Variance Inflation Factor: As a Condition for the Inclusion of Suppressor Variable(s) in Regression Analysis. *Open Journal of Statistics*, *05*(07), 754–767. https://doi.org/10.4236/ojs.2015.57075

Barus, O. P., Happy, J., Pangaribuan, J. J., & Nadjar, F. (2022, September). Liver disease prediction using support vector machine and logistic regression model with combination of PCA and SMOTE. In 2022 1st International Conference on Technology Innovation and Its Applications (ICTIIA) (pp. 1-6). IEEE. https://doi.org/10.1109/ICTIIA54654.2022.9935879

Chawla, N. V, Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. In *Journal of Artificial Intelligence Research* (Vol. 16).

Conti, M. E., Rapa, M., Simone, C., Calabrese, M., Bosco, G., Canepari, S., & Astolfi, M. L. (2024). From land to glass: An integrated approach for quality and traceability assessment of top Italian wines. *Food Control*, *158*. https://doi.org/10.1016/j.foodcont.2023.110226

Cortez, P., Cerdeira, A., Almeida, F., Matos, T., & Reis, J. (2009). Modeling wine preferences by data mining from physicochemical properties. *Decision Support Systems*, *47*(4), 547–553. https://doi.org/10.1016/j.dss.2009.05.016, (Diakses Mei 2025).

Croce, R., Malegori, C., Oliveri, P., Medici, I., Cavaglioni, A., & Rossi, C. (2020). Prediction of quality parameters in straw wine by means of FT-IR spectroscopy combined with multivariate data processing. *Food Chemistry*, *305*. https://doi.org/10.1016/j.foodchem.2019.125512

Dahal, K. R., Dahal, J. N., Banjade, H., & Gaire, S. (2021). Prediction of Wine Quality Using Machine Learning Algorithms. *Open Journal of Statistics*, *11*(02), 278–289. https://doi.org/10.4236/ojs.2021.112015

de Amorim, L. B. V., Cavalcanti, G. D. C., & Cruz, R. M. O. (2022). *The choice of scaling technique matters for classification performance*. https://doi.org/10.1016/j.asoc.2022.109924

Gupta, Y. (2018). Selection of important features and predicting wine quality using machine learning techniques. *Procedia Computer Science*, *125*, 305–312. https://doi.org/10.1016/j.procs.2017.12.041

Khakim, E. N. R., Hermawan, A., & Avianto, D. (2023). IMPLEMENTASI CORRELATION MATRIX PADA KLASIFIKASI DATASET WINE. *JIKO (Jurnal Informatika Dan Komputer)*, *7*(1), 158. https://doi.org/10.26798/jiko.v7i1.771

Kumar, S., Agrawal, K., & Mandan, N. (2020, January 1). Red wine quality prediction using machine learning techniques. *2020 International Conference on Computer Communication and Informatics, ICCCI 2020*. https://doi.org/10.1109/ICCCI48352.2020.9104095

Kurniasari, D., Nurul Hidayah, R., & Khoirun Nisa, R. (2024). CLASSIFICATION MODELS FOR ACADEMIC PERFORMANCE: A COMPARATIVE STUDY OF NAÏVE BAYES AND RANDOM FOREST ALGORITHMS IN ANALYZING UNIVERSITY OF LAMPUNG STUDENT GRADES. *Jurnal Teknik Informatika (JUTIF)*, *5*(5), 1267–1276. https://doi.org/10.52436/1.jutif.2024.5.5.2066

Kurniawan, P., & Widi Nugroho, H. (2024). Implementasi Data Mining Dalam Klasifikasi Tingkat Kesenjangan Kompetensi PNS Menggunakan Metode Naive Bayes. *Technology and Science (BITS)*, *6*(2). https://doi.org/10.47065/bits.v6i2.5641

Mor, N. S., Akiva Schools, B., Tigabo Asras, I., Gal, E., Demasia, T., Tarab, E., Ezekiel, N., Nikapros, O., Semimufar, O., Gladky, E., Karpenko, M., Sason, D., Maslov, D., & Mor, O. (2022). *Wine Quality and Type Prediction from Physicochemical Properties Using Neural Networks for Machine Learning: A Free Software for Winemakers and Customers*.

Olteanu, M., Rossi, F., & Yger, F. (2023). *Meta-survey on outlier and anomaly detection*. https://doi.org/10.1016/j.neucom.2023.126634

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2012). *Scikit-learn: Machine Learning in Python*. http://arxiv.org/abs/1201.0490

Raschka, S. (2018). *Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning*. http://arxiv.org/abs/1811.12808

Romindo, R., Barus, O. P., Pangaribuan, J. J., Pratama, Y. A., & Wiliem, E. (2022). Implementasi Algoritma Support Vector Machine Terhadap Klasifikasi Pose Balet. *Building of Informatics, Technology and Science (BITS)*, *4*(3). https://doi.org/10.47065/bits.v4i3.2647

Sinaga, T., Candra, A., & Purnama, B. (2024). *2024 2nd International Conference On Technology Innovation And Its Applications*. IEEE. https://doi.org/https://doi.org/10.1109/ICTIIA61827.2024.10761639

Vu, D. H., Muttaqi, K. M., & Agalgaonkar, A. P. (2015). A variance inflation factor and backward elimination based robust regression model for forecasting monthly electricity demand using climatic variables. *Applied Energy*, *140*, 385–394. https://doi.org/10.1016/j.apenergy.2014.12.011

Wayan, N., Praditya, P. Y., Kunci, K., & Komputer, J. S. (2023). Prediksi Kualitas Red Wine dan White Wine Menggunakan Data Mining. *JOURNAL SHIFT VOL*, *3*.

Zahedi, L., Mohammadi, F. G., Rezapour, S., Ohland, M. W., & Amini, M. H. (2021). *Search Algorithms for Automated Hyper-Parameter Tuning*. http://arxiv.org/abs/2104.14677