

Classificazione della qualità del vino rosso

Introduzione

Nell'industria vitivinicola, la valutazione della qualità del vino rappresenta un processo critico che tradizionalmente richiede l'intervento di sommelier ed esperti degustatori, con costi elevati e scarsa riproducibilità. La possibilità di classificare automaticamente la qualità del vino a partire dalla sua composizione fisico-chimica consentirebbe ai produttori di vino, alle cantine e ai distributori di standardizzare i controlli qualitativi, ridurre i costi di valutazione e ottimizzare le strategie di pricing. Un modello predittivo basato su analisi di laboratorio può inoltre supportare le decisioni nelle fasi di vinificazione, permettendo interventi correttivi tempestivi.

Obiettivo

Sviluppare un sistema di classificazione della qualità del vino rosso basato sulla sua composizione fisico-chimica. La variabile target è 'WineQuality', che identifica il punteggio di qualità del vino assegnato da esperti sommelier su scala da 3 a 8.

Indicazioni

Il dataset è fornito in due file XLSX (wine_train.xlsx e wine_test.xlsx) contenente sia i parametri fisico-chimici sia il target di classificazione. Il dataset di training può eventualmente essere suddiviso per un validation set. Si ricorda che per una corretta valutazione del modello, i dati di test non devono mai essere utilizzati nella fase di addestramento (training).

Attributi

fixed_acidity: Acidità fissa del vino, principalmente acido tartarico (numerico).

volatile_acidity: Acidità volatile, principalmente acido acetico — livelli elevati causano sapore di aceto (numerico).

citric_acid: Contenuto di acido citrico, aggiunge freschezza e sapore (numerico).

residual_sugar: Zucchero residuo dopo la fermentazione in g/l (numerico).

chlorides: Contenuto di sale nel vino (numerico).

free_sulfur_dioxide: Forma libera di SO₂, previene ossidazione e crescita microbica (numerico).

total_sulfur_dioxide: Quantità totale di SO₂ libera e legata (numerico).

density: Densità del vino, correlata al contenuto di alcol e zucchero (numerico).

pH: Misura dell'acidità o basicità su scala 0–14 (numerico).

sulphates: Additivo che contribuisce ai livelli di SO₂, ha proprietà antimicrobiche e antiossidanti (numerico).

alcohol: Percentuale di alcol in volume (numerico).

WineQuality (label): Variabile target: punteggio di qualità da 3 (pessima) a 8 (eccellente), assegnato da esperti.

3: qualità_molto_scarsa

4: qualità_scarsa

5: qualità_nella_media

6: qualità_buona

7: qualità_molto_buona

8: qualità_eccellente

Acidity_Sulfur_Interaction: Interazione derivata tra acidità fissa e solfati (numerico).

Synthetic_Noise_Flag: Indicatore che indica la presenza rumore nei dati (0 = assente, 1 = presente).

Alcohol_Density_Ratio: Rapporto tra gradazione alcolica e densità (numerico).

Volatile_Enhanced: Indicatore basato sull'acidità volatile elevata (0 = normale, 1 = elevata).