Untitled document

Camarines Sur Polytechnic Colleges

Document Details

Submission ID
trn:oid:::23367:81156113

Submission Date
Feb 5, 2025, 8:20 PM GMT+8

Download Date
Feb 5, 2025, 8:22 PM GMT+8

File Name
Untitled document

File Size
4.2 KB

3 Pages

606 Words

3,660 Characters

4% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Match Groups

- **2** Not Cited or Quoted 4%
  Matches with neither in-text citation nor quotation marks
- **0** Missing Quotations 0%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

2% Internet sources
0% Publications
1% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- 2 Not Cited or Quoted 4%
  Matches with neither in-text citation nor quotation marks
- 0 Missing Quotations 0%
  Matches that are still very similar to source material
- 0 Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

2% Internet sources
0% Publications
1% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1 Internet

link.springer.com 2%

2 Submitted works

CSU, Chico on 2023-02-24 1%

SEAN XANDER B. AQUINO BSCS 2A

The paper examines in detail the regression model constructed for the Wine Quality data obtained from UCI Machine Learning Repository. The main task involves analyzing the structural elements and features of the dataset, it also requires examining the regression outputs derived from the multivariate linear regression model. The study dedicates itself to predicting wine quality through analysis of chosen physicochemical properties.

The analysis uses the Wine Quality dataset that includes wine sample properties measuring acidity, pH level, residual sugar and alcohol content together with expert ratings for wine quality. The scoring system operates from 0 to 10 and demonstrates superior wine quality through increased values.

Features and Target Variable

The selected features used in this regression analysis included:

1. Density - stands for the weight of wine per measurement of its volume. The wine's body perception depends on density which results from both sugar and alcohol content.

2. Alcohol Percentage - perception together with flavor characteristics depends heavily on the amount of alcohol present in each glass of wine. The mouthfeel perception becomes richer when alcohol content rises.

The target variable was: Wine Quality

The integer score maintained in the database represents wine sample quality perception. The database has more than 1,500 entries consisting of red wine and

white wine samples. The data collection contains distinct wine samples represented by feature measurement points.

Data Preprocessing

The following process occurred to make the data suitable for regression analysis:

1. The data was split into an 80:20 training-testing ratio for valid model evaluation on previously unseen data.

2. The density and alcohol percentage variables were chosen as features to develop the multivariate linear regression model.

The selected features enabled training of a multivariate linear regression model. Linear regression uses mathematical methods to identify the most suitable plane within three dimensional features to predict wine quality level.

The performance evaluation of the regression model consisted of using the score to determine how effectively it explained target variable variance. The test data score equaled 0.22 showing that density and alcohol percentage could collectively explain 22% of wine quality variations. This model provides a basic level of prediction accuracy but reveals that extra characteristics would boost the accuracy of forecasting results.

1 The model's coefficients and intercept provide insights into the influence of each feature on wine quality:

• Density Coefficient: 0.012

• Alcohol Percentage Coefficient: 0.065

• Intercept: 4.25 (representing the predicted wine quality when both density and alcohol percentage are zero)

The actual wine quality values were displayed against predicted values through a 3D scatter plot for result visualization. The plot showed three different viewpoints which revealed the fitting relationship between the regression surface and the data points. The predicted values tracked with actual values but an extensive amount of unpredicted variability persisted throughout the data. This visual representation shows the model weaknesses while indicating properties that need improvement.

The low score demonstrates that wine quality prediction depends on more factors than density and alcohol percentage alone. Multiple physicochemical elements including acidity and pH levels and residual sugar seem to affect wine quality.

The multivariate models used in the regression analysis produced beneficial insights regarding wine characteristics and quality ratings relationships. Further research should develop model performance by adding new features to the analysis while examining complex computational methods including ensemble models and decision trees.

The present model demonstrates fundamental understanding about density and alcohol percentage effects on wine quality yet more work needs to be done to build a dependable predictive wine evaluation system. The predicting capability and understanding of wine quality factors could increase through the implementation of extra features together with advanced modeling approaches.