

Collected Experiment Report: DecisionTreeClassifier - RandomForestClassifier - GradientBoostingClassifier .

09/12/2024

Introduction

This is an automated report for the Classification on Red Wine Quality data; the following models have been analyzed:

- DecisionTreeClassifier
- RandomForestClassifier
- GradientBoostingClassifier

Experiment description:

Experiment for classifying the Quality of red wine

Model setup

The models have been used for the following forecast purposes:

- one_step

The models have been optimized using the following hyperparameters:

- max_depth: [6, 8, 10, 15, 20]
- criterion: ['friedman_mse', 'squared_error', 'gini']
- random_state: [42]
- min_samples_split: [5, 10, 50, 150]
- min_samples_leaf: [10, 25, 50, 100]
- scaler: [None, StandardScaler(), MinMaxScaler(), RobustScaler()]

And with the following search algorithms:

- grid
- random

The used performance measure is the accuracy measure.

Dataset setup

The baseline dataset used for these forecasts is
the 'Wine Quality Dataset' dataset: *'Wine Quality Dataset for red wine'*.

The test size used for the forecasts is 0.2.

○ Dataset 1

- name: Full_dataset
- dataset_type: multivariate
- prediction_type: one_step
- components: ['one_step_target', 'feature_columns']

Results

For the models; DecisionTreeClassifier, RandomForestClassifier, GradientBoostingClassifier, the following models and datasets yielded the best results.

The best model for one_step forecasting.

The best model for one_step forecasting is the GradientBoostingClassifier model.

The model has been trained on the Full_dataset dataset.

The best score for the one_step forecasting is 0.540625.

Best GradientBoostingClassifier forecast over time

