Paul Wecker, Robert Wienröder, Vipin Singh

Machine Learning 2 - Project

# Group Members

## Paul Wecker

## Robert Wienröder

## Vipin Singh

# Dataset: Red Wine Quality

## Source:

## Kaggle:

## <https://www.kaggle.com/datasets/uciml/red-wine-quality-cortez-et-al-2009>

## UCI machine learning repository:

## <https://archive.ics.uci.edu/dataset/186/wine+quality>

## Synopsis:

## The dataset contains physiochemical properties of red wine alongside with a quality score (target variable) based on sensory data. The dataset scores are not distributed equally, since there a lot more “normal” quality wines than “poor” or “excellent” quality wines.

## Number of observations: 1599

## Variables:

|  |  |  |
| --- | --- | --- |
| Fixed acidity |  | Continuous |
| Volatile acidity |  | Continuous |
| Citric acid |  | Continuous |
| Residual sugar |  | Continuous |
| Chlorides |  | Continuous |
| Free sulfur dioxide |  | Continuous |
| Total sulfur dioxide |  | Continuous |
| Density |  | Continuous |
| pH |  | Continuous |
| Sulphates |  | Numeric |
| Alcohol |  | Numeric |
| Quality |  | Ordinal |

# Chosen ML-Methods

## Non-linear models (Spline smoothing)

## Pursuit projection regression