jupyter DSA5102 Project Last Checkpoint: 17 hours ago

File Edit View Run Kernel Settings Help

Markdown

1.Preface

In this project, we focus on the dataset `Red Wine Quality` from Kaggle API. I choose my personal interest in wine. After spending weeks tasting wines, I made some guess my learning methods to understand how chemical properties affect wine quality and thier im

[1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

[2]:
```
sns.set_context('notebook', font_scale=1.25, rc={"lines.linewidth": 2.5}
sns.set_style("darkgrid")
np.random.seed(123)  # Tried different seeds - 123 gave most stable resu
```

2.Data Import and Cleaning

This is a dataset from `Kaggle` which contains 1599 records of wine qualities related to physicochemical (inputs) and sensory (output) variables are available (i.e. there is no dat

Input variables (based on physicochemical tests):

- fixed acidity,
- volatile acidity,