

# Wine Quality Prediction System through a Data Mining Approach

Presented by:

**Munshi Saifuzzaman**  
**Reg. No. 2016331094**  
**Dept. of CSE, SUST**

Presented to:

**Md. Forhad Rabbi**  
**Associate Professor**  
**Dept. of CSE, SUST**

on **May 28, 2021**



## » Initials

### Wine

- \* an alcoholic drink typically made from fermented grapes
- \* **Once** viewed as a luxury good
- \* nowadays wine is increasingly enjoyed by a **wider range** of consumers

### Idea Overview

- \* Build an **interface to predict** the **quality** of the **red wine**. (Through data mining approach)
- \* **Result** of the system: Quality of wine, given **chemical information** and **machine learning model**.



## » Motivations

### Wine itself!

- \* Lowers bad cholesterol
- \* Keeps heart healthy
- \* Regulates blood sugar
- \* Reduces the risk of cancer and many more!

### Wine Industry

- \* To **support its growth**, wine industry is investing in new technologies for both **wine making and selling** processes.

## » Motivations (cont.)

### Wine certification

- \* prevents the illegal adulteration of wines (to safeguard human health)
- \* assures quality for the wine market

### Quality Assessment

- \* often part of the certification process
- \* used to improve wine making (by identifying the most influential factors)
- \* used to stratify wines such as premium brands (useful for setting prices)



## » Existing Approach Limitations

## Limitations

Decisions on wine quality prediction are mostly done **scarce** and considers **small datasets**.

- \* **1991's** the "Wine" dataset includes **178 examples** with measurements of 13 chemical constituents
- \* **1997's** the "Wine" dataset includes **170 samples** from Germany but **predict 100% accurately**.
- \* **2001's** wine dataset includes **only 36 examples** were used and 6% error achieved.





## » Contributions

We

- \* **Proposed** a data mining approach to predict human wine taste preferences that is based on easily available analytical tests
- \* Compared to other domain, a **large dataset** is considered with **white and red vinho verde** samples from northwest Portugal have **increased by 36%** from 1997 to 2007

Why Portugal's dataset?

- \* **top ten** wine exporting country, with **3.17%** of the market share in 2005
- \* Exports of its vinho verde wine (from the northwest region) have increased by 36



## » Advantages of this Project

- \* Such model is **useful to support the oenologist** wine tasting evaluations and improve wine production.
- \* similar techniques **can help in target marketing** by modeling consumer tastes from niche markets.

## » Project Demonstration

### Important Files

Actual paper link is given **here**

Dataset Link **here**

### Project

Project Github Link is given **here**

Thanks for your attention!  
Good Day