**Team name: Taste Buds**

**Final Project proposal**

**Team Members:** Abdul Hussain, Gyuheui Choi, Vwerosuo Uzezi



**Project topic:** Wine Quality prediction

**Data Source:** <https://archive.ics.uci.edu/ml/datasets/Wine+Quality>

**Research paper**:

Predicting quality of wine based on chemical attributes Amelia Lemionet, Yi Liu & Zhenxiang Zhou: Stanford University 2009: <http://cs229.stanford.edu/proj2015/245_report.pdf>

**Abstract:**

Wine tasting is the sensory examination of the quality of wine. The practice of wine tasting is almost as ancient as wine production. There are stories recorded of the Sumerian King Gilgamesh in the 3rd Century BC, being able to distinguish between the different beers and wines of the region. A more formalized standard of wine tasting was established in the late 14th century and has continued to evolve till present day. Specialized terminologies are used to describe the various qualities that a wine uniquely possesses such as what region of the world the ingredients were grown, pH and flavors.

The main objective of this project is to predict the best quality of wine based on 12 physiochemical parameters.

Based on the outcome, we ultimately would love to test the predicted outcome against a team of industry wine tasting experts, to see if the results can be validated.

**Expected Results:**

Our result will be a model that can be applied to wine datasets across the wine manufacturing industry. Results will be evaluated based on percentage test error

**Methods**

We will utilize python and different libraries to explore, create model and visualize results.

Other specialized modelling tools like K-Nearest Neighbors, Weighted-Linear Regression and Neural Network will be used for model prediction

Time permitting, we may create a webpage UI where users can input values on the front end and the model will print to screen if the wine taste is good or not.

**Data input fields**

Input variables (based on physicochemical tests):

1 - fixed acidity

2 - volatile acidity

3 - citric acid

4 - residual sugar

5 - chlorides

6 - free sulfur dioxide

7 - total sulfur dioxide

8 - density

9 - pH

10 - sulphates

11 - alcohol

Output variable (based on sensory data):

12 - quality (score between 0 and 10)