



# PREDICTION OF THE CLASSES OF WINES

Agi Tri Uzila



# Table of Contents

## 01. About the Datasets

---

Problem  
Statement, data  
exploration

## 02. Model Development

---

Functions  
definition for the  
model

## 03. Model Training

---

Data  
pre-processing,  
model training

## 04. Model Testing

---

Target Prediction  
using the model

## 05. Model Evaluation

---

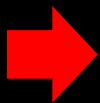
Accuracy  
calculation, cost  
function depiction

# About the Datasets

## Objective

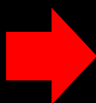
Predict the classes of wines (class 0, class 1, or class 2) from the chemical analysis results of the wines

## 2 Datasets



### Training Dataset

142 entries



### Testing Dataset

36 entries

## 13 Numerical Features

- Alcohol
- Malic acid
- Ash
- Alcalinity of ash
- Magnesium
- Total phenols
- Flavanoids
- Nonflavanoid phenols
- Proanthocyanins
- Color intensity
- Hue
- OD280/OD315 of diluted wines
- Proline

## 1 Categorical Target

- Class 0
- Class 1
- Class 2



**Multiclass  
Classification**

# Model Development

## Logistic Regression

$$h_{\theta}(x) = \frac{1}{1 + e^{(-\theta^T x)}}$$

Sigmoid Function



Cost Function

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m [-y^{(i)} \log(h_{\theta}(x^{(i)}))$$

$$- (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)}))] + \frac{\lambda}{2m} \sum_{j=1}^n \theta_j^2$$



Derivative of Cost Function

$$\frac{\partial J(\theta)}{\partial \theta_j} = \left( \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \right) + \frac{\lambda}{m} \theta_j$$



Gradient Descent Function

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

Variables

$x$  : feature ( $n \times 1$ )

$y$  : target ( $m \times 1$ )

$\theta$  : parameter ( $n \times 1$ )

$\lambda$  : regularization strength

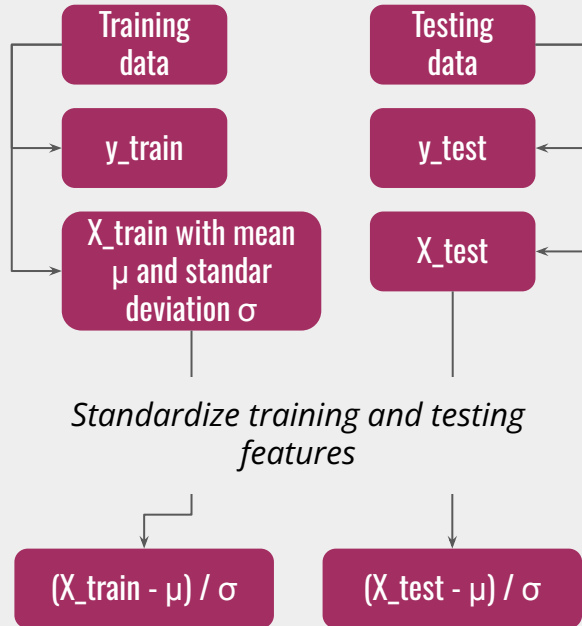
$\alpha$  : learning rate



# Model Training

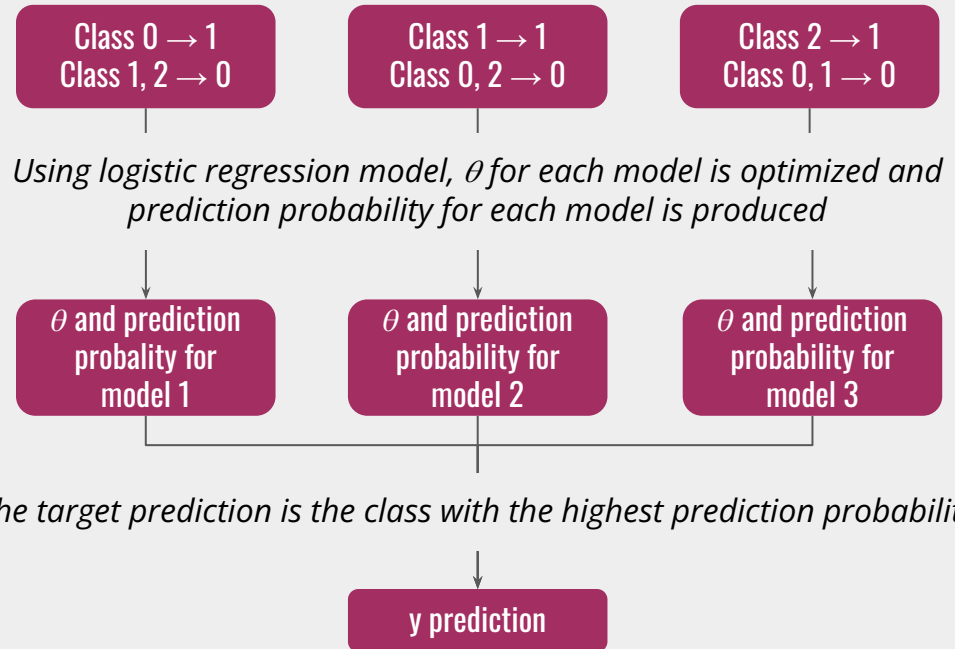
## Data Pre-Processing

*Split the data into features and target*



## Model Training: One-vs-Rest Method

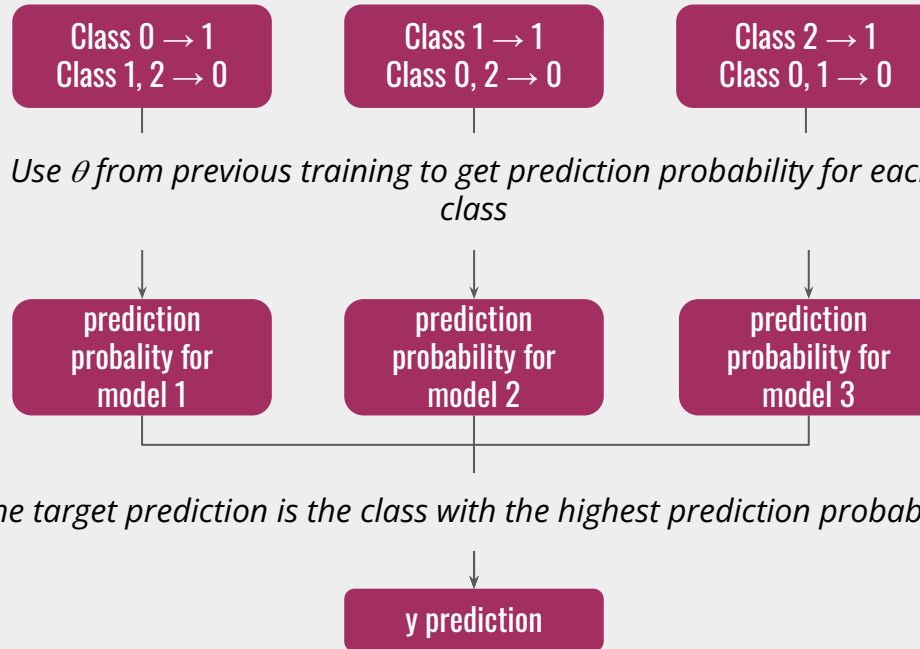
*Build a binary classification model for each class of wines, thus 3 models are constructed*



# Model Testing

## Model Testing: One-vs-Rest Method

*Test the model for each class of wines*



# Model Evaluation

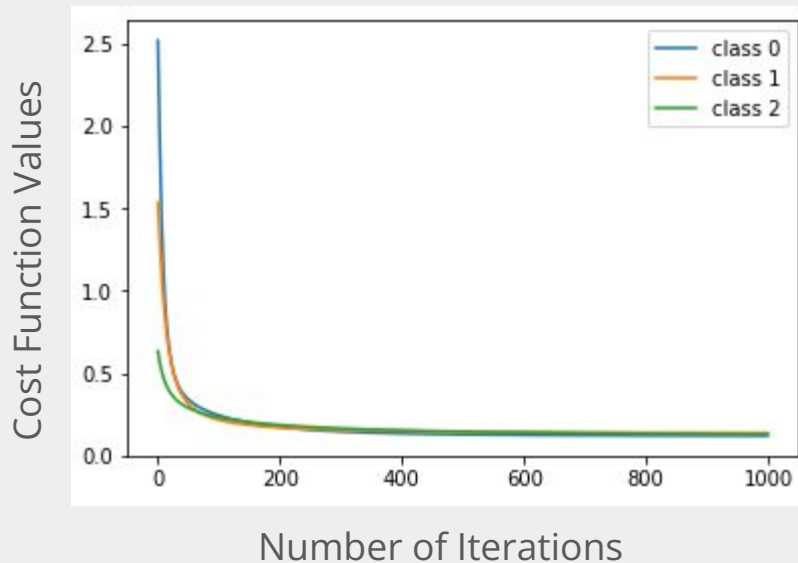
**0.986**

Training  
Accuracy

**0.972**

Testing  
Accuracy

Cost Function Values at Each Iteration





**THANK YOU**

