Using Machine Learning for Milk Grading

1.

import numpy as np # linear algebra

import pandas as pd # data processing, CSV file I/O (e.g. pd.read\_csv)

import matplotlib.pyplot as plt #plotting

%matplotlib inline

plt.rcParams['figure.figsize'] = (15,10) #Set the default figure size

plt.style.use('ggplot') #Set the plotting method

from sklearn.model\_selection import train\_test\_split #Split the data into train and test

from sklearn.ensemble import RandomForestClassifier #Forest for prediction and regression

from sklearn.metrics import mean\_squared\_error #Error testing

from sklearn.metrics import classification\_report #Report of Classification

milk = pd.read\_csv("dataset.csv") #Get the milk dataset

print(milk)

using this commands csv file can be read in jupyter notebook.

2.

Columns

pH, fat-rate, density, temperature

test for different samples with time range 0h-6h.

get the average values of each column.

import numpy as np # linear algebra

import pandas as pd # data processing, CSV file I/O (e.g. pd.read\_csv)

import matplotlib.pyplot as plt #plotting

%matplotlib inline

plt.rcParams['figure.figsize'] = (15,10) #Set the default figure size

plt.style.use('ggplot') #Set the plotting method

from sklearn.model\_selection import train\_test\_split #Split the data into train and test

from sklearn.ensemble import RandomForestClassifier #Forest for prediction and regression

from sklearn.metrics import mean\_squared\_error #Error testing

from sklearn.metrics import classification\_report #Report of Classification

milk = pd.read\_csv("Milk Grading (1).csv") #Get the milk dataset

#Check for Null Values

print(milk)

meanV = np.mean(milk['pH'])

print(meanV)

median

x = numpy.median(speed)

The Mode value is the value that appears the most number of times:

99, 86, 87, 88, 111, 86, 103, 87, 94, 78, 77, 85, 86 = 86

The SciPy module has a method for this.

from scipy import stats  
  
speed = [99,86,87,88,111,86,103,87,94,78,77,85,86]  
  
x = stats.mode(speed)  
  
print(x)