**CS480 – Project Phase 2**

**Assigned on: Monday, 11/2/2015**

**Due: Sunday, 11/15/2015, 11:59pm**

Report Summary 1

class Agent\_sguhatha(Agent) 1

def train(self, X, y) 1

def predict\_prob\_of\_excellent(self, x) 2

ClassSummary 2

ClassMarking 2

Final 2

Mean 2

StDev 2

ProbCalc 2

ClassProbCalc 2

Prediction 2

OUTPUT 3

SIMULATION RESULTS ON dataset1 3

SIMULATION RESULTS ON dataset2 4

SIMULATION RESULTS ON dataset3 5

SIMULATION RESULTS ON dataset4 6

# Report Summary

The Phase 2 of the project tells us to replicated a supervised agent learning and predicting program.

To implement the same I have used Naïve Bayes approach.

My project consist of one file:

* agent\_sguhatha.py

There is one class and two default functions defined:

## class Agent\_sguhatha(Agent)

This is the default class defined under my file to define the training and prediction functions internally.

## def train(self, X, y)

As the name suggest, the function is used to train the agent.

This is basically done by building a classifier by the user of Naïve Bayes approach.

## def predict\_prob\_of\_excellent(self, x)

The function takes the feature to be predicted as the input parameter and returns the probability of the feature being excellent or not.

There are few additional functions defined as well:

## ClassSummary

This function returns a a list based on class labels and its corresponding attributes information. The data consists of the mean and standard deviation of respective attributes.

## ClassMarking

This function is used by ClassSummary to marks the rows of data based on their class labels. The output of this function gives us a list with all the rows mapped to it respective class label.

## Final

It is also used by the ClassSummary function to do the main calculation of the mean and standard deviation based on the data that is passed on my ClassMarking.

## Mean

This calculates mean of all the data of an attribute marked to a class label.

## StDev

This calculates standard deviation of all the data of an attribute marked to a class label.

## ProbCalc

This is the main function used in the prediction of the class label.

This calculates the probability of the features belonging to a excellent or trash nature, given a attribute data, mean and standard deviation.

## ClassProbCalc

This takes the list array of mean and standard deviation of attribute and the test data as its input parameter. It internally calls ProbCalc with individual attribute and its respective mean and standard deviation to get the output of what are the chances of that feature being part of excellent or trash.

## Prediction

Now that the probability of all the attribute is calculated, we find the probability of all the attribute combined being in which of the class label. Finally we send the maximum probability of the class label as the result.

# 

# OUTPUT

**--------------------------------------------------**

## SIMULATION RESULTS ON dataset1

**--------------------------------------------------**

**Wealth (the larger the better)**

**Agent\_fixed\_prob\_0.00: $0.00**

**Agent\_fixed\_prob\_0.25: $806,000.00**

**Agent\_fixed\_prob\_0.50: $1,265,000.00**

**Agent\_fixed\_prob\_0.75: $1,071,000.00**

**Agent\_fixed\_prob\_1.00: $30,000.00**

**Agent\_sguhatha: $1,815,000.00**

**Prediction Log-loss (the smaller the better)**

**Agent\_fixed\_prob\_0.00: 11,582.00**

**Agent\_fixed\_prob\_0.25: 840.28**

**Agent\_fixed\_prob\_0.50: 693.15**

**Agent\_fixed\_prob\_0.75: 833.69**

**Agent\_fixed\_prob\_1.00: 11,443.85**

**Agent\_sguhatha: 3,223.62**

**Prediction Error (the smaller the better)**

**Agent\_fixed\_prob\_0.00: 503**

**Agent\_fixed\_prob\_0.25: 503**

**Agent\_fixed\_prob\_0.50: 497**

**Agent\_fixed\_prob\_0.75: 497**

**Agent\_fixed\_prob\_1.00: 497**

**Agent\_sguhatha: 140**

**--------------------------------------------------**

## SIMULATION RESULTS ON dataset2

**--------------------------------------------------**

**Wealth (the larger the better)**

**Agent\_fixed\_prob\_0.00: $0.00**

**Agent\_fixed\_prob\_0.25: $318,000.00**

**Agent\_fixed\_prob\_0.50: $45,000.00**

**Agent\_fixed\_prob\_0.75: $-637,000.00**

**Agent\_fixed\_prob\_1.00: $-2,410,000.00**

**Agent\_sguhatha: $600,000.00**

**Prediction Log-loss (the smaller the better)**

**Agent\_fixed\_prob\_0.00: 5,963.70**

**Agent\_fixed\_prob\_0.25: 572.22**

**Agent\_fixed\_prob\_0.50: 693.15**

**Agent\_fixed\_prob\_0.75: 1,101.75**

**Agent\_fixed\_prob\_1.00: 17,062.16**

**Agent\_sguhatha: 3,200.59**

**Prediction Error (the smaller the better)**

**Agent\_fixed\_prob\_0.00: 259**

**Agent\_fixed\_prob\_0.25: 259**

**Agent\_fixed\_prob\_0.50: 741**

**Agent\_fixed\_prob\_0.75: 741**

**Agent\_fixed\_prob\_1.00: 741**

**Agent\_sguhatha: 139**

**--------------------------------------------------**

## SIMULATION RESULTS ON dataset3

**--------------------------------------------------**

**Wealth (the larger the better)**

**Agent\_fixed\_prob\_0.00: $0.00**

**Agent\_fixed\_prob\_0.25: $1,278,000.00**

**Agent\_fixed\_prob\_0.50: $2,445,000.00**

**Agent\_fixed\_prob\_0.75: $2,723,000.00**

**Agent\_fixed\_prob\_1.00: $2,390,000.00**

**Agent\_sguhatha: $2,995,000.00**

**Prediction Log-loss (the smaller the better)**

**Agent\_fixed\_prob\_0.00: 17,016.10**

**Agent\_fixed\_prob\_0.25: 1,099.56**

**Agent\_fixed\_prob\_0.50: 693.15**

**Agent\_fixed\_prob\_0.75: 574.42**

**Agent\_fixed\_prob\_1.00: 6,009.75**

**Agent\_sguhatha: 3,223.62**

**Prediction Error (the smaller the better)**

**Agent\_fixed\_prob\_0.00: 739**

**Agent\_fixed\_prob\_0.25: 739**

**Agent\_fixed\_prob\_0.50: 261**

**Agent\_fixed\_prob\_0.75: 261**

**Agent\_fixed\_prob\_1.00: 261**

**Agent\_sguhatha: 140**

**--------------------------------------------------**

## SIMULATION RESULTS ON dataset4

**--------------------------------------------------**

**Wealth (the larger the better)**

**Agent\_fixed\_prob\_0.00: $0.00**

**Agent\_fixed\_prob\_0.25: $748,000.00**

**Agent\_fixed\_prob\_0.50: $1,120,000.00**

**Agent\_fixed\_prob\_0.75: $868,000.00**

**Agent\_fixed\_prob\_1.00: $-260,000.00**

**Agent\_sguhatha: $1,635,000.00**

**Prediction Log-loss (the smaller the better)**

**Agent\_fixed\_prob\_0.00: 10,914.25**

**Agent\_fixed\_prob\_0.25: 808.42**

**Agent\_fixed\_prob\_0.50: 693.15**

**Agent\_fixed\_prob\_0.75: 865.55**

**Agent\_fixed\_prob\_1.00: 12,111.60**

**Agent\_sguhatha: 3,384.80**

**Prediction Error (the smaller the better)**

**Agent\_fixed\_prob\_0.00: 474**

**Agent\_fixed\_prob\_0.25: 474**

**Agent\_fixed\_prob\_0.50: 526**

**Agent\_fixed\_prob\_0.75: 526**

**Agent\_fixed\_prob\_1.00: 526**

**Agent\_sguhatha: 147**