BINDIYA C.M., 4MTI7CS028

PAGE NO: 11

DATE : 22/12/20

EXP. NO. : 5

5.	Write a program to implement the naive Bayesian classifier for a sample training dataset stored as a
	classifier for a sample training dataset stored as a
	. CSV file. Compute the accuracy of the classifier,
	considering few test data sets
	emport csv, random, math
	emport statestes as st
	def loadcsv (filename):
	lines = csv. reader (open (felename, "r"))
	dataset = list (19nex)
	for i in range (len (dataset)):
	dataset[i] = [float(x) for x in dataset[i]]
	return dataset
	The state of the s
	def splitDataset (dataset, splitRatio):
	test Size = int (len (dataset) * split Ratio);
	trainset = list (dataset);
	testset = []
	while len(test set) < test Size:
	endex = random. randrange (len (train set)); test set. append (train set, pop (endex))
	test set. append (trainset, pop (index))
	return [trainset, test Set]
	def separate By Class (dataset): separated = ()
	separated = ()
	for in range (len (dataget)):
	separated = 1 )  for i in range (len (dataset)):  x = dataset[i]
	if (x[-i] not in separated):  ARUN'S

PAGE NO: 12

DATE:
EXP. NO.:

-	
	Separated [x [-1]] = []
	separated [x[-i]]. append(x)
	return separated
	det compute-mean std (dataset):
	mean_std=[(st. mean (attribute), st. stder (attribute))
	for attribute in zip (* dataset)];
	del mean_std[-1]
	return mean_std
	def Summarize By Class (dataset):
	separated = separate Byclau (datoret);
	Summary = 1
	for classvalue, instances in separated items():
10.33	summary [class Value] = compute_mean_std (Pytance)
	return summary
	def estimate Probability (x, mean, stder):
	exponent = math. exp(- (math. pow (x-mean, 2)/
A STATE OF THE PARTY OF THE PAR	(2 x math. pow (stdev, 2))))
	return (1/(math. sgrt (2 x math. pr) * stdev)) * exponent
	def colculate Claw Probabilities (summaries, test vector):
	P= { }
	for class Value, classemmaries in summaries. Hem():
	p[clasvalue]=1
	for i in range (len (claus Summarter)): mean, stder = claus Summarter [i]iruns
	mean, stder = clau Summaria [1] irunis

BINDIYA C.M., 4MTI7CS028

PAGE NO: 13
DATE:
EXP. NO.:

```
X = test Vector [1]
                 p[class Value] * = estimate Probability (x, mean,
                                                         stder);
    return p
det predict (summarter, test vector):
     all-p = Galculate Class Probabilities (summarce, test Vector)
     bettlabel, bett Prob = None, -1
     for Ibl, p in all-p. stems():
          if bet Label & None or p> bet Prob;
             beltProb = p
             bethabel = 161
     return bettlabel
det perform_claur fecation (summarlel, teltset):
     prediction = [ ]
     for in range (lon(textset)):
result = predict (summarrer, textset[1])
          prediction. append (result)
      return predictions
def getAccuracy (testset, prediction):
correct = 0
       for i en range (len(textset)):
           ef testset[][-i] == prediction[i]:
               correct += 1
        return (correct/floatlen(testSet))) x 100.0
                                                      ARUNS
```

BINDIYA C.M., 4MTI7CSO28

PAGE NO: 14
DATE:
EXP. NO::

datalet = loadcsv('dia	beter. cs	1);	1/ EASTER	1日 图 图 1			
print ('Pima Indian Diabetes Dataset loaded')							
print ('Total instances available: ', len (dataset))							
prent ('Total attributes present: ', len (dataset [0]-1)							
print ("First Five instances of dataset: ")							
for 9 en range (5):		And a B	1030	sept also			
print (i+1, ':', dataget [i])							
splet Rate = 0.2							
frainingset, test set = split Dataset (dataset, split Ratio)							
print ('in Dataset is split into training and testing set')							
print ('Training examples = log In Testing examples = 113'.							
	format (len (training set), len (test set)))						
Summariel = Summarize By Class (training Set);							
prediction = perform_claufteation (summarier, tellSet)							
accuracy - setAccuracy (testSet, Drediction)							
pront ('in Accuracy of the Native Bayesian classifier is: ',							
accuracy)			150130	Herman			
A VOCAS	310 12 B	HISO 27	College				
1,000	00.1	1001	20.1	6			
	919-10	V10 - 0	(C) (1)				
The Laboratory of the Control of the	E arab	/ room	01:0				
A CONTRACTOR OF THE PARTY OF TH	atr.b.			A DESCRIPTION OF			
	TITLE TO	14.0	1110				
	18800	1100	781.0 1.50	b, b dillocat			
				ARUN'S —			

## Output:-

Pema Indian Diabetes Datalet loaded

Total postances avaslable: 768

Total attributes present: 8

FIRE Five intences of dataset:

1: [6.0, 148.0, 72.0, 35.0, 0.0, 33.6, 0.627, 50,0, 1.0]

2:[1.0,85.0,66.0,29.0,0.0,26.6,0.351,31.0,0.0]

3:[8.0,183.0,64.0,0.0,0.0,23.3,0.672,32.0,1.0]

4:[1.0, 89.0, 66.0, 23.0, 94.0, 28.1, 0.167, 21.0, 0.0]

5: [0.0,137.0,40.0,35.0,168.0,43.1,2.288,33.0,1.0]

Dataset & split into training and testing set.

Training example = 615

Accuracy of the Naive Bayescan Classifier is: 75.16339869281046

Consider the traditional addition to the control of the control of

and Coppel Dison