Assignment 2

1. Nominal attributes are categorical which take pre-defined value that have finite set of possibilities. The values of the nominal specification only serve as category labels or names. Numeric attributes measure numbers, either real or integer valued that are continuous. **In this case, use the nominal specification.** Bayes classifier works with categorical variables because it is unlikely to find exact matches for numerical variables.

2.

handinfan

Count of party	Column Labels		
Row Labels	democrat	republican	Grand Total
n	50	83	133
y	71	23	94
Grand Total	121	106	227
n	50/121	83/106	
y	71/121	23/106	

water

Count of party	Column Labels		
Row Labels	democrat	republican	Grand Total
n	67	57	124
y	54	49	103
Grand Total	121	106	227
n	67/121	57/106	
y	54/121	49/106	

budget

Count of party	Column Labels		
Row Labels	democrat	republican	Grand Total
n	18	89	107
y	103	17	120
Grand Total	121	106	227
n	18/121	89/106	
y	103/121	17/106	

physfee

Count of party	Column Labels		
Row Labels	democrat	republican	Grand Total
n	116	1	117
y	5	105	110
Grand Total	121	106	227
n	116/121	1/106	
y	5/121	105/106	

salvador

Count of party	Column		
Row Labels	democra		
n			
y			
Grand Total	121	106	227
n	97/121	5/106	
y	24/121	101/106	

		republican	Grand Total
n	68	14	82
y	53	92	145
Grand Total	121	106	227
n		06	
y		06	

sattest

Count of party	Column Labels		
Row Labels	democrat	republican	Grand Total
n	28	77	105
y	93	29	122
Grand Total	121	106	227
n	28/121	77/106	
y	93/121	29/106	

contras

Count of			
Row Labels	democrat	republican	Grand Total
n	20	90	110
y	101	16	117
Grand Total	121	106	227
n	20/121	90/106	
y	101/121	16/106	

missile

Count of party	Column Labels		
Row Labels	democrat	republican	Grand Total
n	25	91	116
y	96	15	111
Grand Total	121	106	227
n	25/121	91/106	
y	96/121	15/106	

imm

Count of party	Column Labels		
Row Labels	democrat	republican	Grand Total
n	55	44	99
y	66	62	128
Grand Total	121	106	227
n	55/121	44/106	
y	66/121	62/106	

synfuel

Count of party	Column Labels		
Row Labels	democrat	republican	Grand Total
n	60	89	149
y	61	17	78
Grand Total	121	106	227
n	60/121	89/106	
y	61/121	17/106	

edu

Count of party	Column Labels		
Row Labels	democrat	republican	Grand Total
n	106	15	121
y	15	91	106
Grand Total	121	106	227
n	106/121	15/106	
y	15/121	91/106	

Assignment Project Exam Help
https://eduassistpro.github.io/
Add WeChat edu_assist_pro

rttosue

Count of party	Column Labels		
Row Labels	democrat	republican	Grand Total
n	86	17	103
y	35	89	124
Grand Total	121	106	227
n	86/121	17/106	
y	35/121	89/106	

crime

Count of party	Column Labels		
Row Labels	democrat	republican	Grand Total
n	80	2	82
y	41	104	145
Grand Total	121	106	227
n	80/121	2/106	
y	41/121	104/106	

export

Count of party	Column Labels		
Row Labels	democrat	republican	Grand Total
n	49	94	143
y	72	12	84
Grand Total	121	106	227
n	49/121	94/106	
y	72/121	12/106	

exportsa

Count of party	Column Labels		
Row Labels	democrat	republican	Grand Total
n	7	36	43
y	114	70	184
Grand Total	121	106	227
n	7/121	36/106	
y	114/121	70/106	

party

Row Labels	Count of party
democrat	121
republican	106
Grand Total	227
democrat	121/227
republican	106/227

Test Case	Actual data	Classification result
1	Democrat	Democrat
2	Republican	Republican
3		blican
4		blican
5		ocrat

Assignment Project Exam Help
https://eduassistpro.github.io/
Add WeChat edu_assist_pro

Based on the results summarized in the table a f accurate classification from the classifier is 4/5 = 80%

party, handinfan, water, budget, physfee, salvador, relinsch, sattest, contras, missile, imm, synfuel, edu, rttosue, crime, export, exportsa

Test Case 1: democrat, y, n, y, n, n, n, y, y, y, n, n, n, n, n, y, y
5 cases: 2 points

For democrat = 71/121 * 67/121 * 103/121 * 116/121 * 97/121 * 68/121 * 93/121 * 101/121 * 96/121 * 55/121 * 60/121 * 106/121 * 86/121 * 80/121 * 72/121 * 114/121 * 121/227
= 1.68587 x 10E-03

1.68587 x 10E-03 / (1.68587 x 10E-03 + 3.34840 x 10E-14) **= 100%**

For republican = 23/106 * 57/106 * 17/106 * 1/106 * 5/106 * 14/106 * 29/106 * 16/106 * 15/106 * 44/106 * 89/106 * 15/106 * 17/106 * 2/106 * 12/106 * 70/106 * 106/227
= 3.34840 x 10E-14

3.34840 x 10E-14 / (1.68587 x 10E-03 + 3.34840 x 10E-14) **= 0%**

Classification result is accurate; test case predicts democrat. Based on training data, result is democrat.

Test Case 2: republican, n, y, n, y, y, y, n, n, n, n, n, y, y, y, n, y

For democrat = 50/121 * 54/121 * 18/121 * 5/121 * 24/121 * 53/121 * 28/121 * 20/121 * 25/121 * 55/121 * 60/121 * 15/121 * 35/121 * 41/121 * 49/121 * 114/121 * 121/227
= 4.33484 x 10E-10

4.33484 x 10E-10 / (4.33484 x 10E-10 + 8.88499 x 10E-03) **= 0%**

For republican = 83/106 * 49/106 * 89/106 * 105/106 * 101/106 * 92/106 * 77/106 * 90/106 * 91/106 * 44/106 * 89/106 * 91/106 * 89/106 * 104/106 * 94/106 * 70/106 * 106/227
= 8.88499 x 10E-03

8.88499 x 10E-03 / (4.33484 x 10E-10 + 8.88499 x 10E-03) **= 100%**

Classification result is accurate; test case predicts republican. Based on training data, result is republican.

Assignment Project Exam Help

Test Case 3: r y, n, y

https://eduassistpro.github.io/

For democrat = 5 * 28/121 * 20/121 * 25/121 * 55/121 * 60/121 * 106/121 * 35/121 * 4 21 * 121/227
= 3.06329 x 10E-09

Add WeChat edu_assist_pro

3.06329 x 10E-09 / (3.06329 x 10E-09 + 1.46456 x 10E-03) **= 0%**

For republican = 83/106 * 49/106 * 89/106 * 105/106 * 101/106 * 92/106 * 77/106 * 90/106 * 91/106 * 44/106 * 89/106 * 15/106 * 89/106 * 104/106 * 94/106 * 70/106 * 106/227
= 1.46456 x 10E-03

1.46456 x 10E-03 / (3.06329 x 10E-09 + 1.46456 x 10E-03) **= 100%**

Classification result is accurate; test case predicts republican. Based on training data, result is republican.

Test Case 4: democrat, y, y, y, y, y, y, n, n, n, n, y, y, y, y, n, y

For democrat = 71/121 * 54/121 * 103/121 * 5/121 * 24/121 * 53/121 * 28/121 * 20/121 * 25/121 * 55/121 * 61/121 * 15/121 * 35/121 * 41/121 * 49/121 * 114/121 * 121/227
= 3.58100 x 10E-09

3.58100 x 10E-09 / (3.58100 x 10E-09 + 8.98306 x 10E-05) **= 0%**

For republican = 23/106 * 49/106 * 17/106 * 105/106 * 101/106 * 92/106 * 77/106 * 90/106 * 91/106 * 44/106 * 17/106 * 91/106 * 89/106 * 104/106 * 94/106 * 70/106 * 106/227
= 8.98306 x 10E-05

8.98306 x 10E-05 / (3.58100 x 10E-09 + 8.98306 x 10E-05) **= 100%**

Classification result is not accurate; test case predicts democrat. Based on training data, result is republican.

Test Case 5: democrat, n, y, y, n, n, y, y, y, y, n, y, n, n, y, y, y

For democrat = 50/121 * 54/121 * 103/121 * 116/121 * 97/121 * 53/121 * 93/121 * 101/121 * 96/121 * 55/121 * 61/121 * 106/121 * 86/121 * 41/121 * 72/121 * 114/121 * 121/227
= 3.88592 x 10E-04

3.88592 x 10E-04 / (3.88592 x 10E-04 + 6.78002 x 10E-12) **= 100%**

Assignment Project Exam Help

For republican = 86/106 * 49/106 * 17/106 * 11/106 * 5/106 * 92/106 * 29/106 * 16/106 * 15/106 * 44/106 * 17/106 * 15/106 * 17/106 * 104/106 * 12/106 * 70/106 * 106/227
= 6.78002 x 10E-

6.78002 x 10E-12 https://eduassistpro.github.io/

Classification result is accurate; test case predi n training data, result is democrat.

Add WeChat edu_assist_pro

Confusion matrix for training-no-NULL.ARFF wit o-NULL.ARFF:

a	b		<-- classified as
2	0	|	a = republican
1	2	|	b = democrat

3.

(a) Based on this confusion matrix, estimate the overall accuracy of the classifier.
921 + 28 + 17 + 374 = 1340
921 + 374 = 1295
17 + 28 = 45
1295/1340 = 96.6%

(b) Estimate the stratified accuracies of the classifier.
For a: 921/(921+28) = **97.0%**
For b: 374/(374+17) = **95.7%**