



20IT6302: MACHINE LEARNING HOME ASSIGNMENT-3 QUESTIONS

s. n o	QUESTION	CO	BTL																																																																																																																			
1	<table><tr><th>User ID</th><th>Gender</th><th>Age</th><th>EstimatedSalary</th><th>Purchased</th></tr><tr><td>15624510</td><td>Male</td><td>19</td><td>19000</td><td>0</td></tr><tr><td>15810944</td><td>Male</td><td>35</td><td>20000</td><td>0</td></tr><tr><td>15668575</td><td>Female</td><td>26</td><td>43000</td><td>0</td></tr><tr><td>15603246</td><td>Female</td><td>27</td><td>57000</td><td>0</td></tr><tr><td>15804002</td><td>Male</td><td>19</td><td>76000</td><td>0</td></tr><tr><td>15728773</td><td>Male</td><td>27</td><td>58000</td><td>0</td></tr><tr><td>15598044</td><td>Female</td><td>27</td><td>84000</td><td>0</td></tr><tr><td>15694829</td><td>Female</td><td>32</td><td>150000</td><td>1</td></tr><tr><td>15600575</td><td>Male</td><td>25</td><td>33000</td><td>0</td></tr><tr><td>15727311</td><td>Female</td><td>35</td><td>65000</td><td>0</td></tr><tr><td>15570769</td><td>Female</td><td>26</td><td>80000</td><td>0</td></tr><tr><td>15606274</td><td>Female</td><td>26</td><td>52000</td><td>0</td></tr><tr><td>15746139</td><td>Male</td><td>20</td><td>86000</td><td>0</td></tr><tr><td>15704987</td><td>Male</td><td>32</td><td>18000</td><td>0</td></tr><tr><td>15628972</td><td>Male</td><td>18</td><td>82000</td><td>0</td></tr><tr><td>15697686</td><td>Male</td><td>29</td><td>80000</td><td>0</td></tr><tr><td>15733883</td><td>Male</td><td>47</td><td>25000</td><td>1</td></tr><tr><td>15617482</td><td>Male</td><td>45</td><td>26000</td><td>1</td></tr><tr><td>15704583</td><td>Male</td><td>46</td><td>28000</td><td>1</td></tr><tr><td>15621083</td><td>Female</td><td>48</td><td>29000</td><td>1</td></tr><tr><td>15649487</td><td>Male</td><td>45</td><td>22000</td><td>1</td></tr><tr><td>15736760</td><td>Female</td><td>47</td><td>49000</td><td>1</td></tr></table> <p>Check whetehr the customer with features Gender=Male; Age=26; estimated salary=60000 will purchase a car or not.</p>	User ID	Gender	Age	EstimatedSalary	Purchased	15624510	Male	19	19000	0	15810944	Male	35	20000	0	15668575	Female	26	43000	0	15603246	Female	27	57000	0	15804002	Male	19	76000	0	15728773	Male	27	58000	0	15598044	Female	27	84000	0	15694829	Female	32	150000	1	15600575	Male	25	33000	0	15727311	Female	35	65000	0	15570769	Female	26	80000	0	15606274	Female	26	52000	0	15746139	Male	20	86000	0	15704987	Male	32	18000	0	15628972	Male	18	82000	0	15697686	Male	29	80000	0	15733883	Male	47	25000	1	15617482	Male	45	26000	1	15704583	Male	46	28000	1	15621083	Female	48	29000	1	15649487	Male	45	22000	1	15736760	Female	47	49000	1	CO2	Apply
User ID	Gender	Age	EstimatedSalary	Purchased																																																																																																																		
15624510	Male	19	19000	0																																																																																																																		
15810944	Male	35	20000	0																																																																																																																		
15668575	Female	26	43000	0																																																																																																																		
15603246	Female	27	57000	0																																																																																																																		
15804002	Male	19	76000	0																																																																																																																		
15728773	Male	27	58000	0																																																																																																																		
15598044	Female	27	84000	0																																																																																																																		
15694829	Female	32	150000	1																																																																																																																		
15600575	Male	25	33000	0																																																																																																																		
15727311	Female	35	65000	0																																																																																																																		
15570769	Female	26	80000	0																																																																																																																		
15606274	Female	26	52000	0																																																																																																																		
15746139	Male	20	86000	0																																																																																																																		
15704987	Male	32	18000	0																																																																																																																		
15628972	Male	18	82000	0																																																																																																																		
15697686	Male	29	80000	0																																																																																																																		
15733883	Male	47	25000	1																																																																																																																		
15617482	Male	45	26000	1																																																																																																																		
15704583	Male	46	28000	1																																																																																																																		
15621083	Female	48	29000	1																																																																																																																		
15649487	Male	45	22000	1																																																																																																																		
15736760	Female	47	49000	1																																																																																																																		
2	<table><tr><th>Height (in cms)</th><th>Weight (in kgs)</th><th>T Shirt Size</th></tr><tr><td>158</td><td>58</td><td>M</td></tr><tr><td>158</td><td>59</td><td>M</td></tr><tr><td>158</td><td>63</td><td>M</td></tr><tr><td>160</td><td>59</td><td>M</td></tr><tr><td>160</td><td>60</td><td>M</td></tr><tr><td>163</td><td>60</td><td>M</td></tr><tr><td>163</td><td>61</td><td>M</td></tr><tr><td>160</td><td>64</td><td>L</td></tr><tr><td>163</td><td>64</td><td>L</td></tr><tr><td>165</td><td>61</td><td>L</td></tr><tr><td>165</td><td>62</td><td>L</td></tr><tr><td>165</td><td>65</td><td>L</td></tr><tr><td>168</td><td>62</td><td>L</td></tr><tr><td>168</td><td>63</td><td>L</td></tr></table> <p>Suppose we have height, weight and T-shirt size of some customers. Predict the T-shirt size of a new customer 'Monica' has height 161cm and weight 61kg.</p>	Height (in cms)	Weight (in kgs)	T Shirt Size	158	58	M	158	59	M	158	63	M	160	59	M	160	60	M	163	60	M	163	61	M	160	64	L	163	64	L	165	61	L	165	62	L	165	65	L	168	62	L	168	63	L	CO2	Apply																																																																						
Height (in cms)	Weight (in kgs)	T Shirt Size																																																																																																																				
158	58	M																																																																																																																				
158	59	M																																																																																																																				
158	63	M																																																																																																																				
160	59	M																																																																																																																				
160	60	M																																																																																																																				
163	60	M																																																																																																																				
163	61	M																																																																																																																				
160	64	L																																																																																																																				
163	64	L																																																																																																																				
165	61	L																																																																																																																				
165	62	L																																																																																																																				
165	65	L																																																																																																																				
168	62	L																																																																																																																				
168	63	L																																																																																																																				

3	<table border="1"> <thead> <tr> <th>ID</th><th>SPEED</th><th>AGILITY</th><th>DRAFT</th></tr> </thead> <tbody> <tr><td>1</td><td>2.50</td><td>6.00</td><td>yes</td></tr> <tr><td>2</td><td>3.75</td><td>8.00</td><td>no</td></tr> <tr><td>3</td><td>2.25</td><td>5.50</td><td>yes</td></tr> <tr><td>4</td><td>3.25</td><td>8.25</td><td>no</td></tr> <tr><td>5</td><td>2.75</td><td>7.50</td><td>no</td></tr> <tr><td>6</td><td>4.50</td><td>5.00</td><td>yes</td></tr> <tr><td>7</td><td>3.50</td><td>5.25</td><td>yes</td></tr> <tr><td>8</td><td>3.00</td><td>3.25</td><td>no</td></tr> <tr><td>9</td><td>4.00</td><td>4.00</td><td>no</td></tr> <tr><td>10</td><td>4.25</td><td>3.75</td><td>yes</td></tr> </tbody> </table> <p>The speed and agility of the 10 college athletes along with whether they have drafted or not is given in the above table. Predict whether an athlete with speed=6.75 and agility=3.00 will be drafted or not.</p>	ID	SPEED	AGILITY	DRAFT	1	2.50	6.00	yes	2	3.75	8.00	no	3	2.25	5.50	yes	4	3.25	8.25	no	5	2.75	7.50	no	6	4.50	5.00	yes	7	3.50	5.25	yes	8	3.00	3.25	no	9	4.00	4.00	no	10	4.25	3.75	yes	CO2	Apply
ID	SPEED	AGILITY	DRAFT																																												
1	2.50	6.00	yes																																												
2	3.75	8.00	no																																												
3	2.25	5.50	yes																																												
4	3.25	8.25	no																																												
5	2.75	7.50	no																																												
6	4.50	5.00	yes																																												
7	3.50	5.25	yes																																												
8	3.00	3.25	no																																												
9	4.00	4.00	no																																												
10	4.25	3.75	yes																																												
4	<table border="1"> <thead> <tr> <th>ID</th><th>SPEED</th><th>AGILITY</th><th>DRAFT</th></tr> </thead> <tbody> <tr><td>11</td><td>2.00</td><td>2.00</td><td>no</td></tr> <tr><td>12</td><td>5.00</td><td>2.50</td><td>no</td></tr> <tr><td>13</td><td>8.25</td><td>8.50</td><td>no</td></tr> <tr><td>14</td><td>5.75</td><td>8.75</td><td>yes</td></tr> <tr><td>15</td><td>4.75</td><td>6.25</td><td>yes</td></tr> <tr><td>16</td><td>5.50</td><td>6.75</td><td>no</td></tr> <tr><td>17</td><td>5.25</td><td>9.50</td><td>yes</td></tr> <tr><td>18</td><td>7.00</td><td>4.25</td><td>no</td></tr> <tr><td>19</td><td>7.50</td><td>8.00</td><td>yes</td></tr> <tr><td>20</td><td>7.25</td><td>5.75</td><td>yes</td></tr> </tbody> </table> <p>The speed and agility of the 10 college athletes along with whether they have drafted or not is given in the above table. Predict whether an athlete with speed=6.75 and agility=3.00 will be drafted or not.</p>	ID	SPEED	AGILITY	DRAFT	11	2.00	2.00	no	12	5.00	2.50	no	13	8.25	8.50	no	14	5.75	8.75	yes	15	4.75	6.25	yes	16	5.50	6.75	no	17	5.25	9.50	yes	18	7.00	4.25	no	19	7.50	8.00	yes	20	7.25	5.75	yes	CO2	Apply
ID	SPEED	AGILITY	DRAFT																																												
11	2.00	2.00	no																																												
12	5.00	2.50	no																																												
13	8.25	8.50	no																																												
14	5.75	8.75	yes																																												
15	4.75	6.25	yes																																												
16	5.50	6.75	no																																												
17	5.25	9.50	yes																																												
18	7.00	4.25	no																																												
19	7.50	8.00	yes																																												
20	7.25	5.75	yes																																												
5	<p>Cluster the following eight points (with (x, y) representing locations) into three clusters using k-means clustering:</p> <p>A1(2, 10), A2(2, 5), A3(8, 4), A4(5, 8), A5(7, 5), A6(6, 4), A7(1, 2), A8(4, 9)</p>	CO2	Apply																																												
6	<p>Cluster the following ten points (with (x, y) representing locations) into two clusters using k-means clustering:</p> <p>A1 (2,10), A2 (2,6), A3 (11,11), A4 (6,9), A5 (6,4), A6 (1,2), A7 (5,10), A8 (4,9), A9 (10,12), A10 (7,5)</p>	CO2	Apply																																												
7	<p>Cluster the following eight points (with (x, y) representing locations) into two clusters using k-means clustering:</p> <p>A1(2, 10), A2(2, 5), A3(8, 4), A4(5, 8), A5(7, 5), A6(6, 4), A7(1, 2), A8(4, 9)</p>	CO2	Apply																																												
8	<p>Cluster the following ten points (with (x, y) representing locations) into three clusters using k-means clustering:</p> <p>A1 (2,10), A2 (2,6), A3 (11,11), A4 (6,9), A5 (6,4), A6 (1,2), A7 (5,10), A8 (4,9), A9 (10,12), A10 (7,5)</p>	CO2	Apply																																												

9	<table><tr><td></td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr><tr><td>1</td><td>0</td><td></td><td></td><td></td><td></td></tr><tr><td>2</td><td>9</td><td>0</td><td></td><td></td><td></td></tr><tr><td>3</td><td>3</td><td>7</td><td>0</td><td></td><td></td></tr><tr><td>4</td><td>6</td><td>5</td><td>9</td><td>0</td><td></td></tr><tr><td>5</td><td>11</td><td>10</td><td>2</td><td>8</td><td>0</td></tr></table> <p>Draw the dendrogram by Performing the agglomerative clustering and find the best number of clusters for the 5 objects with the above distance matrix.</p>		1	2	3	4	5	1	0					2	9	0				3	3	7	0			4	6	5	9	0		5	11	10	2	8	0	CO2	Apply																											
	1	2	3	4	5																																																													
1	0																																																																	
2	9	0																																																																
3	3	7	0																																																															
4	6	5	9	0																																																														
5	11	10	2	8	0																																																													
10	<table><tr><td></td><td>a</td><td>b</td><td></td><td>P1</td><td>P2</td><td>P3</td><td>P4</td><td>P5</td></tr><tr><td>Point</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>P1</td><td>0.07</td><td>0.83</td><td>P1</td><td>0</td><td></td><td></td><td></td><td></td></tr><tr><td>P2</td><td>0.85</td><td>0.14</td><td>P2</td><td>1.04139</td><td>0</td><td></td><td></td><td></td></tr><tr><td>P3</td><td>0.66</td><td>0.89</td><td>P3</td><td>0.59304</td><td>0.77369</td><td>0</td><td></td><td></td></tr><tr><td>P4</td><td>0.49</td><td>0.64</td><td>P4</td><td>0.46098</td><td>0.61612</td><td>0.30232</td><td>0</td><td></td></tr><tr><td>P5</td><td>0.80</td><td>0.46</td><td>P5</td><td>0.81841</td><td>0.32388</td><td>0.45222</td><td>0.35847</td><td>0</td></tr></table> <p>Draw the dendrogram by Performing the agglomerative clustering and find the best number of clusters for the 5 objects with the above distance matrix.</p>		a	b		P1	P2	P3	P4	P5	Point									P1	0.07	0.83	P1	0					P2	0.85	0.14	P2	1.04139	0				P3	0.66	0.89	P3	0.59304	0.77369	0			P4	0.49	0.64	P4	0.46098	0.61612	0.30232	0		P5	0.80	0.46	P5	0.81841	0.32388	0.45222	0.35847	0	CO2	Apply
	a	b		P1	P2	P3	P4	P5																																																										
Point																																																																		
P1	0.07	0.83	P1	0																																																														
P2	0.85	0.14	P2	1.04139	0																																																													
P3	0.66	0.89	P3	0.59304	0.77369	0																																																												
P4	0.49	0.64	P4	0.46098	0.61612	0.30232	0																																																											
P5	0.80	0.46	P5	0.81841	0.32388	0.45222	0.35847	0																																																										
11	<table><tr><td>Distance</td><td>a</td><td>b</td><td>c</td><td>d</td><td>e</td></tr><tr><td>a</td><td>0</td><td>2</td><td>6</td><td>10</td><td>9</td></tr><tr><td>b</td><td>2</td><td>0</td><td>5</td><td>9</td><td>8</td></tr><tr><td>c</td><td>6</td><td>5</td><td>0</td><td>4</td><td>5</td></tr><tr><td>d</td><td>10</td><td>9</td><td>4</td><td>0</td><td>3</td></tr><tr><td>e</td><td>9</td><td>8</td><td>5</td><td>3</td><td>0</td></tr></table> <p>Draw the dendrogram by Performing the agglomerative clustering and find the best number of clusters for the 5 objects with the above distance matrix.</p>	Distance	a	b	c	d	e	a	0	2	6	10	9	b	2	0	5	9	8	c	6	5	0	4	5	d	10	9	4	0	3	e	9	8	5	3	0	CO2	Apply																											
Distance	a	b	c	d	e																																																													
a	0	2	6	10	9																																																													
b	2	0	5	9	8																																																													
c	6	5	0	4	5																																																													
d	10	9	4	0	3																																																													
e	9	8	5	3	0																																																													
12	<table><tr><td></td><td>p1</td><td>p2</td><td>p3</td><td>p4</td><td>p5</td><td>p6</td></tr><tr><td>p1</td><td>0.0000</td><td>0.2357</td><td>0.2218</td><td>0.3688</td><td>0.3421</td><td>0.2347</td></tr><tr><td>p2</td><td>0.2357</td><td>0.0000</td><td>0.1483</td><td>0.2042</td><td>0.1388</td><td>0.2540</td></tr><tr><td>p3</td><td>0.2218</td><td>0.1483</td><td>0.0000</td><td>0.1513</td><td>0.2843</td><td>0.1100</td></tr><tr><td>p4</td><td>0.3688</td><td>0.2042</td><td>0.1513</td><td>0.0000</td><td>0.2932</td><td>0.2216</td></tr><tr><td>p5</td><td>0.3421</td><td>0.1388</td><td>0.2843</td><td>0.2932</td><td>0.0000</td><td>0.3921</td></tr><tr><td>p6</td><td>0.2347</td><td>0.2540</td><td>0.1100</td><td>0.2216</td><td>0.3921</td><td>0.0000</td></tr></table> <p>Draw the dendrogram by Performing the agglomerative clustering and find the best number of clusters for the 5 objects with the above distance matrix.</p>		p1	p2	p3	p4	p5	p6	p1	0.0000	0.2357	0.2218	0.3688	0.3421	0.2347	p2	0.2357	0.0000	0.1483	0.2042	0.1388	0.2540	p3	0.2218	0.1483	0.0000	0.1513	0.2843	0.1100	p4	0.3688	0.2042	0.1513	0.0000	0.2932	0.2216	p5	0.3421	0.1388	0.2843	0.2932	0.0000	0.3921	p6	0.2347	0.2540	0.1100	0.2216	0.3921	0.0000	CO2	Apply														
	p1	p2	p3	p4	p5	p6																																																												
p1	0.0000	0.2357	0.2218	0.3688	0.3421	0.2347																																																												
p2	0.2357	0.0000	0.1483	0.2042	0.1388	0.2540																																																												
p3	0.2218	0.1483	0.0000	0.1513	0.2843	0.1100																																																												
p4	0.3688	0.2042	0.1513	0.0000	0.2932	0.2216																																																												
p5	0.3421	0.1388	0.2843	0.2932	0.0000	0.3921																																																												
p6	0.2347	0.2540	0.1100	0.2216	0.3921	0.0000																																																												

13

Name	Blood Type	Give Birth	Can Fly	Live in Water	Class
human	warm	yes	no	no	mammals
python	cold	no	no	no	reptiles
salmon	cold	no	no	yes	fishes
whale	warm	yes	no	yes	mammals
frog	cold	no	no	sometimes	amphibians
komodo	cold	no	no	no	reptiles
bat	warm	yes	yes	no	mammals
pigeon	warm	no	yes	no	birds
cat	warm	yes	no	no	mammals
leopard shark	cold	yes	no	yes	fishes
turtle	cold	no	no	sometimes	reptiles
penguin	warm	no	no	sometimes	birds
porcupine	warm	yes	no	no	mammals
eel	cold	no	no	yes	fishes
salamander	cold	no	no	sometimes	amphibians
gila monster	cold	no	no	no	reptiles
platypus	warm	no	no	no	mammals
owl	warm	no	yes	no	birds
dolphin	warm	yes	no	yes	mammals
eagle	warm	no	yes	no	birds

Find the class of species if {blood type=cold, give birth=yes, can fly=yes, live in water=sometimes}

CO3

Apply

14

1-Consider the following training dataset for a binary classification problem

Object number	Home owner	Marital status	Sex	Income	Defaulted borrower
1	Yes	Married	Female	150	No
2	Yes	Married	Male	220	Yes
3	Yes	Divorced	Female	75	No
4	No	Single	Female	80	Yes
5	Yes	Single	Male	110	No
6	No	Divorced	Male	65	Yes
7	Yes	Single	Female	90	Yes
8	No	Married	Female	55	No
9	No	Divorced	Male	85	No
10	Yes	Married	Male	95	No

G. Predict a class for the following record using Naive Bayes Classifier

Object number	Home owner	Marital status	Sex	Income	Defaulted borrower
1	Yes	Single	Female	80	??

CO3

Apply

15

Using naïve Bayesian classification find the class label of the tuple:
presbyopic, hypermetrope, yes, normal

Record ID	Age	Spectacle prescription	Astigmatic	Tear production Rate	Class label Lenses
1	Young	Myope	No	Reduced	Noncontact
2	Young	Myope	No	Normal	Soft contact
3	Young	Myope	Yes	Reduced	Noncontact
4	Young	Myope	Yes	Normal	Hard contact
5	Young	Hypermetrope	No	Reduced	Noncontact
6	Young	Hypermetrope	No	Normal	Soft contact
7	Young	Hypermetrope	Yes	Reduced	Noncontact
8	Young	Hypermetrope	Yes	Normal	Hard contact
9	Pre-presbyopic	Myope	No	Reduced	Noncontact
10	Pre-presbyopic	Myope	No	Normal	Soft contact
11	Pre-presbyopic	Myope	Yes	Reduced	Noncontact
12	Pre-presbyopic	Myope	Yes	Normal	Hard contact
13	Pre-presbyopic	Hypermetrope	No	Reduced	Noncontact
14	Pre-presbyopic	Hypermetrope	No	Normal	Soft contact
15	Pre-presbyopic	Hypermetrope	Yes	Reduced	Noncontact
16	Pre-presbyopic	Hypermetrope	Yes	Normal	Noncontact
17	Presbyopic	Myope	No	Reduced	Noncontact
18	Presbyopic	Myope	No	Normal	Noncontact
19	Presbyopic	Myope	Yes	Reduced	Noncontact
20	Presbyopic	Myope	Yes	Normal	Hard contact

CO3

Apply

16

Given a data tuple having the values “*systems*,” “26 . . . 30,” and “46–50K” for the attributes *department*, *age*, and *salary*, respectively, what would a naïve Bayesian classification of the *status* for the tuple be?

The following table consists of training data from an employee database. The data have been generalized. For example, “31 . . . 35” for *age* represents the age range of 31 to 35. For a given row entry, *count* represents the number of data tuples having the values for *department*, *status*, *age*, and *salary* given in that row.

<i>department</i>	<i>status</i>	<i>age</i>	<i>salary</i>	<i>count</i>
sales	senior	31 . . . 35	46K . . . 50K	30
sales	junior	26 . . . 30	26K . . . 30K	40
sales	junior	31 . . . 35	31K . . . 35K	40
systems	junior	21 . . . 25	46K . . . 50K	20
systems	senior	31 . . . 35	66K . . . 70K	5
systems	junior	26 . . . 30	46K . . . 50K	3
systems	senior	41 . . . 45	66K . . . 70K	3
marketing	senior	36 . . . 40	46K . . . 50K	10
marketing	junior	31 . . . 35	41K . . . 45K	4
secretary	senior	46 . . . 50	36K . . . 40K	4
secretary	junior	26 . . . 30	26K . . . 30K	6

Let *status* be the class label attribute.

CO3

Apply

17

Use Naïve Bayesian classification to predict label of the 8th tuple.

Example	Colour	Toughness	Fungus	Appearance	Poisonous
1	Green	Hard	N	Smooth	N
2	Green	Hard	Y	Smooth	N
3	Brown	Soft	N	Wrinkled	N
4	Orange	Hard	N	Wrinkled	Y
5	Green	Soft	Y	Smooth	Y
6	Green	Hard	Y	Wrinkled	Y
7	Orange	Hard	N	Wrinkled	Y
8	Green	Soft	Y	Wrinkled	?

CO3

Apply

18	X Variables				Y Variable	CO3	Apply
	Blood Pressure	Fever	Diabetes	Vomit	Suffering from disease 'Z'		
high	high	yes	no	no			
high	high	yes	yes	no			
low	high	yes	no	yes			
normal	mild	yes	no	yes			
normal	no fever	no	no	yes			
normal	no fever	no	yes	no			
low	no fever	no	yes	yes			
high	mild	yes	no	no			
high	no fever	no	no	yes			
normal	mild	no	no	yes			
high	mild	no	yes	yes			
low	mild	yes	yes	yes			
low	high	no	no	yes			
normal	mild	yes	yes	no			
Low no fever yes yes ?							
Use naïve Bayesian classification for the above label prediction.							
19	Example No.	Color	Type	Origin	Stolen?	CO3	Apply
	1	Red	Sports	Domestic	Yes		
2	Red	Sports	Domestic	No			
3	Red	Sports	Domestic	Yes			
4	Yellow	Sports	Domestic	No			
5	Yellow	Sports	Imported	Yes			
6	Yellow	SUV	Imported	No			
7	Yellow	SUV	Imported	Yes			
8	Yellow	SUV	Domestic	No			
9	Red	SUV	Imported	No			
10	Red	Sports	Imported	Yes			
Use Naïve Bayesian classification for predicting the class label of the instance {color=red, Type=SUV, origin=domestic}							
20	PlayTennis: training examples					CO3	Apply
	Day	Outlook	Temperature	Humidity	Wind	PlayTennis	
D1	Sunny	Hot	High	Weak	No		
D2	Sunny	Hot	High	Strong	No		
D3	Overcast	Hot	High	Weak	Yes		
D4	Rain	Mild	High	Weak	Yes		
D5	Rain	Cool	Normal	Weak	Yes		
D6	Rain	Cool	Normal	Strong	No		
D7	Overcast	Cool	Normal	Strong	Yes		
D8	Sunny	Mild	High	Weak	No		
D9	Sunny	Cool	Normal	Weak	Yes		
D10	Rain	Mild	Normal	Weak	Yes		
D11	Sunny	Mild	Normal	Strong	Yes		
D12	Overcast	Mild	High	Strong	Yes		
D13	Overcast	Hot	Normal	Weak	Yes		
D14	Rain	Mild	High	Strong	No		
Use Naïve Bayesian classification for predicting the class label of the instance {outlook=rainy, temarature=cool, humidity= high, wind=strong.}							

21

Given all the previous patients I've seen (below are their symptoms and diagnosis)...

chills	runny nose	headache	fever	flu?
Y	N	Mild	Y	N
Y	Y	No	N	Y
Y	N	Strong	Y	Y
N	Y	Mild	Y	Y
N	N	No	N	N
N	Y	Strong	Y	Y
N	Y	Strong	N	N
Y	Y	Mild	Y	Y

Do I believe that a patient with the following symptoms has the flu?

chills	runny nose	headache	fever	flu?
Y	N	Mild	Y	?

CO3

Apply

22

Apply the **Naïve Bayes classifier** for the following *Pie dataset* in order to determine the class of the following object.

$x = [\text{crust-size=thick, crust-shade=gray, filling-size=thin, filling shade=white, shape=square}]$

Pie Dataset						
Example	Crust		Filling		Shape	Class
	Size	Shade	Size	Shade		
Ex1	Thick	Gray	Thick	Dark	circle	positive
Ex2	Thick	White	Thick	Dark	circle	positive
Ex3	Thick	Dark	Thick	Gray	triangle	positive
Ex4	Thin	White	Thin	Dark	circle	positive
Ex5	Thick	Dark	Thin	White	square	positive
Ex6	Thick	White	Thin	Dark	circle	positive
Ex7	Thick	Gray	Thick	White	circle	negative
Ex8	Thick	White	Thick	Gray	square	negative
Ex9	Thin	Gray	Thin	Dark	triangle	negative
Ex10	Thick	Dark	Thick	White	circle	negative
Ex11	Thick	White	Thick	Dark	square	negative
Ex12	Thick	White	Thick	Gray	triangle	negative

CO3

Apply

23

Consider the training examples shown in the following table for a binary classification. The table shows a training set for a problem of predicting whether a loan applicant will repay his/her loan obligation or defaulting on his/her loan.

Tid	Home Owner	Marital Status	Annual Income	Defaulted Borrower
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

Using the **NB** approach that we discussed in the class, predict the class label for this test example, $X = (\text{Home Owner} = \text{No}, \text{Marital Status} = \text{Married}, \text{Income} = \$120\text{K})$.

CO3

Apply