**Question-1**

\*\*\* p.s: there was typo in this question in midterm and hence everyone will be given points for this

Consider binary classification where the classes are: C1 and C2. These two classes are equally likely (same probability).

Each data point has 2 dimensional features: X1 and X2 .

The following tables shows the probability distribution across the the classes are: X1 and X2.  .

**Table:1** shows P(C1 | X1)

|  |  |
| --- | --- |
| X1 | P(X1 | C1) |
| -1 | 0.2 |
| 0 | 0.4 |
| 1 | 0.4 |

**Table 2 shows** P(C2 | X1 )

|  |  |
| --- | --- |
| X1 | P(X1 | C2) |
| -1 | 0.3 |
| 0 | 0.6 |
| 1 | 0.1 |

**Table 3 shows** P(C1 | X1)

|  |  |
| --- | --- |
| X2 | P(X2 | C1) |
| -1 | 0.4 |
| 0 | 0.5 |
| 1 | 0.1 |

**Table 4 shows** P(C2 | X2)

|  |  |
| --- | --- |
| X2 | P(X2 | C1) |
| -1 | 0.1 |
| 0 | 0.3 |
| 1 | 0.6 |

We have 2 classes and both of them are equally likely, so P(C1) = ½ =0.5

What will be posterior-probability of P(C1| X1 = −1, X2 = 1)?

=

=0.1

What will be posterior-probability of P(C2| X1 = −1, X2 = 1)= 0.9

What will be class predicted by NB? C2

**Question-2**

Consider the following data set comprised of two input attributes (GPA, Studied) and one output attribute (Outcome):

|  |  |  |
| --- | --- | --- |
| GPA | Studied | Outcome |
| Low | False | Failed |
| Low | True | Passed |
| Medium | False | Failed |
| Medium | True | Passed |
| High | False | Passed |
| High | True | Passed |

What will be entropy for the decision Column (Outcome)?

Entropy(Outcome) = - { (2/6)log2(2/6) + (4/6)log2(4/6)} = 0.92

Entropy(GPA, Outcome) = - [ {(1/3) \*( 1/2log2(1/2) + 1/2log2(1/2))} +  
 {(1/3) \*( 1/2log2(1/2) + 1/2log2(1/2))} + {(1/3) \*( 1log2(1))} ]

= 0.66

 Entropy(Studied, Outcome) = - [ {(1/2) \*( 1/3log2(1/3) + 2/3log2(2/3))} +  
 {(1/2) \*( 1log2(1))} ]

= 0.46

So higher gain from Studied

So root Studied

**Question-3**

After your yearly checkup, the doctor has bad news and good news. The bad news is that you tested positive for a serious disease and that the test is 99% accurate (i.e., the probability of testing positive when you do have the disease is 0.99, as is the probability of testing negative when you don’t have the disease). The good news is that this is a rare disease, striking only 1 in 10,000 people of your age. What is the probability that that you actually have the disease? (answer should contain 2 digits after the decimal point)

We were given:

P(T|D) = 0.99

P(D) = 0.0001

We want:

P(D|T)   
= P(T|D) \* P(D) / P(T)

= ( P(T|D) P(D) ) / ( P(T|D) P(D) + (P(T|!D) P(!D) )

= 0.99 \* 0.0001 / ( 0.99 \* 0.0001 + 0.01 \* 0.9999)

= 0.0098

= 0.01

**Question-4**

Suppose that the following logical sentences are all True and they are in your knowledge base.

Is B True?

Text

Description automatically generated

If B is True, then C is True, & F is True

If C^F is True then B is False. \*\* Contradiction\*\*\*  
So B is False

Is E True?

If B is False E must be True

Chart, bubble chart

Description automatically generated

Q1. What is the state sequence (including A and G) of the optimal solution?

A-B-C-D-G / ABCDG

Q2. What is the total cost of the solution by the breadth-first search (BFS) algorithm (please fill in an integer)?

Hint: The total cost is the sum of the costs along the edges the solution passes through.

10

Q3. Let the heuristic h(state) be  r \* the minimum number of edges the state has to travel to reach G.

For example, h(B) is r\*3 , h(E) is r\*2, h(G) is  r\*0, and h(F) is r\*1.

If r=5, what is the solution by A\* search (please fill in a sequence of states; e.g., AG)  
A-F-G

If r=5, what is the total cost of the solution by A\* search (please fill in an integer)?

9

If there is a tie in choosing which node to expand first (note that, a node in a search algorithm is a sequence of states), you will choose the node with the least number of states.

Q4. Is the heuristic (with r=5 ) admissible (Yes or No)?

No, as it overestimates the distance from a state to the nearest goal state.

Chart, bubble chart

Description automatically generated

The above figure shows a state space graph, in which state A is the start state, and state G is the goal state. Each edge (i.e., action) is directed and associated with a cost.

Q1. What is the state sequence (including A and G) of the optimal solution? [optimal]

Example: AG is the state sequence of a solution, but it may not be optimal.

A-B-C-D-G

Q2. What is the total cost of the solution by the breadth-first search (BFS) algorithm (please fill in an integer)?

Hint: The total cost is the sum of the costs along the edges the solution passes through.

10

Q3. Let the heuristic h(state) be  r \* the minimum number of edges the state has to travel to reach G.

For example, h(B) is r\*3 , h(E) is r\*2, h(G) is  r\*0, and h(F) is r\*1.

If r = 3, what is the solution by A\* search (please fill in a sequence of states; e.g., AG)?

A-F-G

If r = 3, what is the total cost of the solution by A\* search (please fill in an integer)?

9

If there is a tie in choosing which node to expand first (note that, a node in a search algorithm is a sequence of states), you will choose the node with the least number of states.

Q4. Is the heuristic (with r = 3) admissible (Yes or No)?

No

Graphical user interface, text

Description automatically generated with medium confidence

(A)

Graphical user interface, text

Description automatically generated

(B)

**True/False**

1. An agent function takes the action (through the actuator) as the input and outputs the percept (through the sensor). (False, it’s the opposite)
2. Being rational means maximizing the expected utility. (True)
3. If a search problem has solutions and one of the solutions has a finite depth, then BFS is guaranteed to find the solution. (True)
4. A search **problem** with a finite number of states always has a search **tree** of a finite depth. (False, there can be cycles)
5. The sequence of nodes a search algorithm **expands** in order to reach the goal state(s) is called the **solution**. (False, we may need to explore many other states)
6. A heuristic whose value is 0 at every state is an admissible heuristic. (True, as it will never overestimates the distance from a state to the nearest goal state.)
7. If a search problem has solutions and the search tree has a finite depth, then depth-first search (DFS) is **guaranteed** to find the **optimal** solution. (False, found solution may not be optimal)