Name: Enrolment No:	

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
End Semester Examination, May 2022

Programme Name: B. Tech. ME | **Semester : VIII**
Course Name : Artificial Intelligence | **Time : 03 Hrs**
Course Code : CSEG3005P | **Max. Marks : 100**
Nos. of page(s) : 07
Instructions:

SECTION A [5x4]

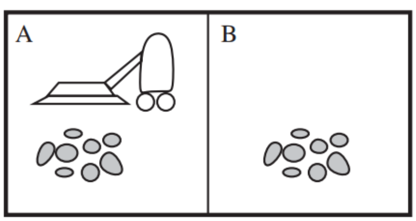
S. No.		Marks	CO
Q 1.	**State True/ False:** *Problem solving is a process of generating solutions from observed data. [T/F]* *A state is defined by the specification of the values of all attributes of interest in the world. [T/F]* *The goal state is the partial description of the solution. [T/F]* *The notation Ǝ !xP(x) states "There exists a unique x such that P(x) is true". [T/F]*	[1+1+1 +1]	CO1
Q 2.	Enlist the four essential properties of any Search Algorithms	4	CO2
Q 3. (a) (b)	Let P(x) denote the statement “x > 10″. What are the truth values of P(11) and P(5)? Let R(x,y) denote the statement “x = y + 1“. What is the truth value of the propositions R (1,3) and R(2,1)?	2+2	CO1
Q 4.	Differentiate between Supervised learning and Un-supervised learning in AI.	4	CO3

<table>
<tr><td>Q5.</td><td>What is the probability that a patient has diseases meningitis with a stiff neck?

Given Data:

A doctor is aware that disease meningitis causes a patient to have a stiff neck, and it occurs 80% of the time. He is also aware of some more facts, which are given as follows:

The Known probability that a patient has meningitis disease is 1/30,000.
The Known probability that a patient has a stiff neck is 2%.

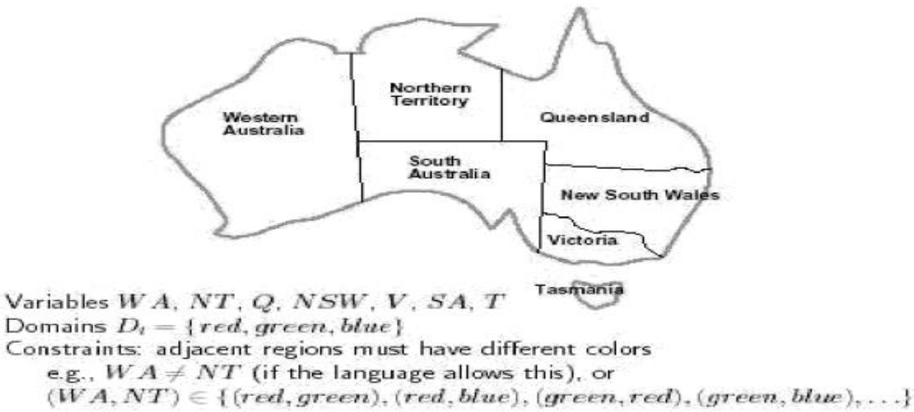
Apply Bayes' theorem to figure out the solution.</td><td>4</td><td>CO3</td></tr>
<tr><td colspan="4">SECTION B [4x10]</td></tr>
<tr><td>Q 6.</td><td>The Vacuum-cleaner world shown in figure below. This particular world has just two locations: squares A and B. The vacuum agent perceives which square it is in and whether there is dirt in the square. It can choose to move left, move right, suck up the dirt, or do nothing.
One very simple agent function is the following: if the current square is dirty, then suck, otherwise move to the other square.


<table><tr><th>Percept Sequence</th><th>Action</th></tr><tr><td>[A, Clean]</td><td>Right</td></tr><tr><td>[A, Dirty]</td><td>Suck</td></tr><tr><td>[B, Clean]</td><td>Left</td></tr><tr><td>[B, Dirty]</td><td>Suck</td></tr><tr><td>[A, Clean], [A, Clean]</td><td>Right</td></tr><tr><td>[A, Clean], [A, Dirty]</td><td>Suck</td></tr></table>
Write a REFLEX-VACCUM-AGENT program
(piece of code)that is invoked for each new percept (location, status) and returns an action each time</td><td>10</td><td>CO2</td></tr>
<tr><td>Q 7.</td><td>Elaborate the performance of Non-parametric models w.r.t. the following:
(a) Performance
(b) Flexibility
(c) Little to no assumptions

Also, discuss the limitations of the non-parametric models in context of the undermentioned:
(a) Overfitting
(b) Speed
(c) Training Data</td><td>10</td><td>CO3</td></tr>
<tr><td>Q 8.</td><td>You are given the task of coloring each region red, green, or blue in such a way that the neighboring regions must not have the same color.

To formulate this as Constraint Satisfaction Problem, we define the variable to be the regions: WA, NT, Q, NSW, V, SA, and T.
The domain of each variable is the set {red, green, blue}.The constraints require</td><td>10</td><td>CO1</td></tr>
</table>

neighboring regions to have distinct colors: for example, the allowable combinations for WA and NT are the pairs {(red,green),(red,blue),(green,red),(green,blue),(blue,red),(blue,green)}.

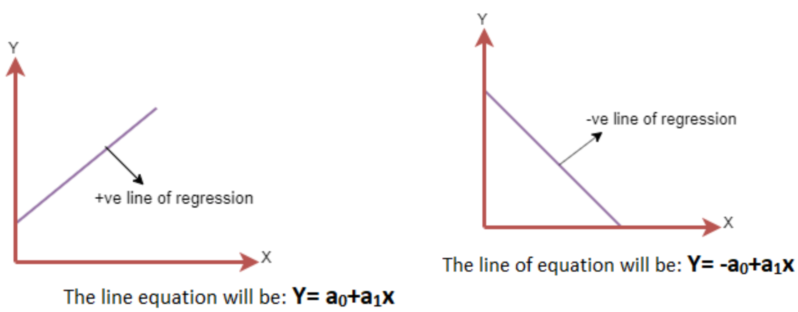


Solve this CSP problem via indicating the color combinations (as provided) satisfying all the constraints suggested in the question.

Q 9.

Linear Regression Line

A linear line showing the relationship between the dependent and independent variables is called a **regression line**. A regression line can show two types of relationship:



Discuss all the important **Assumptions** of Linear Regression. These are some formal checks while building a Linear Regression model, which ensures to get the best possible result from the given dataset.

OR

Suggest the output(values of Estimated Coefficients) of the code applied on Simple Linear Regression:

```
import numpy as np
import matplotlib.pyplot as plt

def estimate_coef(x, y):
    # number of observations/points
    n = np.size(x)
```

10

CO2

```
    # mean of x and y vector
    m_x = np.mean(x)
    m_y = np.mean(y)

    # calculating cross-deviation and deviation about x
    SS_xy = np.sum(y*x) - n*m_y*m_x
    SS_xx = np.sum(x*x) - n*m_x*m_x

    # calculating regression coefficients
    b_1 = SS_xy / SS_xx
    b_0 = m_y - b_1*m_x

    return (b_0, b_1)

def plot_regression_line(x, y, b):
    # plotting the actual points as scatter plot
    plt.scatter(x, y, color = "m",
                marker = "o", s = 30)

    # predicted response vector
    y_pred = b[0] + b[1]*x

    # plotting the regression line
    plt.plot(x, y_pred, color = "g")

    # putting labels
    plt.xlabel('x')
    plt.ylabel('y')

    # function to show plot
    plt.show()

def main():
    # observations / data
    x = np.array([0, 1, 2, 3, 4, 5, 6, 7, 8, 9])
    y = np.array([1, 3, 2, 5, 7, 8, 8, 9, 10, 12])

    # estimating coefficients
    b = estimate_coef(x, y)
    print("Estimated coefficients:\nb_0 = {} \
        \nb_1 = {}".format(b[0], b[1]))

    # plotting regression line
    plot_regression_line(x, y, b)

if __name__ == "__main__":
    main()
```

SECTION-C [20x2]

Q 10.

(a) Consider a fictional dataset that describes the weather conditions for playing a game of golf. Given the weather conditions, each tuple classifies the conditions as fit ("Yes") or unfit ("No") for playing golf.

Here is a tabular representation of our dataset.

	Outlook	Temperature	Humidity	Windy	Play Golf
0	Rainy	Hot	High	False	No
1	Rainy	Hot	High	True	No
2	Overcast	Hot	High	False	Yes
3	Sunny	Mild	High	False	Yes
4	Sunny	Cool	Normal	False	Yes
5	Sunny	Cool	Normal	True	No
6	Overcast	Cool	Normal	True	Yes
7	Rainy	Mild	High	False	No
8	Rainy	Cool	Normal	False	Yes
9	Sunny	Mild	Normal	False	Yes

12

CO3

10	Rainy	Mild	Normal	True	Yes
11	Overcast	Mild	High	True	Yes
12	Overcast	Hot	Normal	False	Yes
13	Sunny	Mild	High	True	No

The dataset is divided into two parts, namely, **feature matrix** and the **response vector**.

- Feature matrix contains all the vectors (rows) of dataset in which each vector consists of the value of **dependent features**. In above dataset, features are ‘Outlook’, ‘Temperature’, ‘Humidity’ and ‘Windy’.
- Response vector contains the value of **class variable** (prediction or output) for each row of feature matrix. In above dataset, the class variable name is ‘Play golf’.

Assumption:

The fundamental Naive Bayes assumption is that each feature makes an:

- independent
- equal

-contribution to the outcome.

With relation to our dataset, this concept can be understood as:

We assume that no pair of features are dependent. For example, the temperature being ‘Hot’ has nothing to do with the humidity or the outlook being ‘Rainy’ has no effect on the winds. Hence, the features are assumed to be independent.

Secondly, each feature is given the same weight (or importance). For example, knowing only temperature and humidity alone can’t predict the outcome accurately. None of the attributes is irrelevant and assumed to be contributing equally to the outcome.

Now, it is time to put a naive assumption to the Bayes’ theorem, which is, independence among the features. So now, you split evidence into the independent parts as shown below and complete the table with your expected computations:

Outlook

	Yes	No	P(yes)	P(no)
Sunny	[redacted]	[redacted]	[redacted]	[redacted]
Overcast	[redacted]	[redacted]	[redacted]	[redacted]
Rainy	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]

Temperature

	Yes	No	P(yes)	P(no)
Hot	[redacted]	[redacted]	[redacted]	[redacted]
Mild	[redacted]	[redacted]	[redacted]	[redacted]
Cool	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]

Humidity

	Yes	No	P(yes)	P(no)
High	[redacted]	[redacted]	[redacted]	[redacted]
Normal	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]

Wind

	Yes	No	P(yes)	P(no)
False	[redacted]	[redacted]	[redacted]	[redacted]
True	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]

Play	P(Yes)/P(No)	
Yes	[redacted]	[redacted]
No	[redacted]	[redacted]
Total	[redacted]	[redacted]

(b)	How do you approach the Strengths and Weakness of **Decision Tree**? **OR** Discuss the **N-Queen problem** in AI and what are the possible methods to solve the same?	8	CO2
Q11. (A)	**Illustrate the following 'Use Cases' of Natural Language Processing in Business:** (i) Sentiment Analysis (ii) Language Translation (iii) Text Extraction (iv) Chat-bots	10+10	CO3

	(v) Topic Classification		
(B)	**Define the following:** (i) Speech Recognition and AI (ii) Information Retrieval (iii) Information Extraction (iv) Statistical Learning (v) Ontological Engineering		