# Solution

## Question 1

### Solution:

Step 1: Calculating Prior Probabilities

|  |  |
| --- | --- |
| **P(C)** | **Value** |
| P(NO) | 3/10 |
| P(MayBe) | 3/10 |
| P(Yes) | 4/10 |

Step 2: Calculating Likelihood

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Student** | | | | | |
| NO | | Maybe | | Yes | |
| P(is\_a\_student|No) | 1/3 | P(is\_a\_student|Maybe) | 3/3 | P(is\_a\_student|Yes) | 2/4 |
| P(not\_a\_student|No) | 2/3 | P(not\_a\_student|Maybe) | 0/3 | P(not\_a\_student|Yes | 2/4 |
| **Salary** | | | | | |
| P(>20k|No) | 1/3 | P(>20k|Maybe) | 2/3 | P(>20k|Yes) | 4/4 |
| P(<20k|No) | 2/3 | P(<20k|Maybe) | 1/3 | P(<20k|Yes) | 0/4 |
| **Studies** | | | | | |
| P(CS|No) | 2/3 | P(CS|Maybe) | 0/3 | P(CS|Yes) | 2/4 |
| P(Math|No) | 1/3 | P(Math|Maybe) | 2/3 | P(Math|Yes) | 1/4 |
| P(Eng|No) | 0/2 | P(Eng|Maybe) | 1/3 | P(Eng|Yes) | 1/4 |
| **Credit Rating** | | | | | |
| P(fair|No) | 1/3 | P(fair|Maybe) | 1/3 | P(fair|Yes) | 2/4 |
| P(Good|No) | 1/3 | P(Good|Maybe) | 1/3 | P(Good|Yes) | 1/4 |
| P(Exc|No) | 1/3 | P(Exc|Maybe) | 1/3 | P(Exc|Yes) | 1/4 |

Step 3: Now calculating for unseen example:

**X --> Student** = Is\_a\_student , **Salary** < 20K, **Studies** = Engineering, **Credit rating** =Good

* P(No|X)=(1/3)\*(1/3)\*(0/3)\*(1/3)\*(3/10)=0
* P(Maybe|X)=(3/3)\*(1/3)\*(1/3)\*(1/3)\*(3/10)=**0.0111**
* P(Yes|X)=(2/4)\*(0/4)\*(1/4)\*(1/4)\*(4/10)=0

As P(Maybe|X) is greater than both P(No|X) and P(Yes|X). So, the class label for given example is ***Maybe***

## **Question 2**

1. **Sentence 1:**

* "Engaging, exciting, and stimulating education can be enriching and empowering, while boring and tedious education can be stultifying and disengaging."
* **Positive words:** Engaging, stimulating, enriching, empowering (4 words)
* **Negative words:** Boring, tedious, stultifying, disengaging (4 words)
* **Polarity of Sentence** = (Number of positive words) - (Number of negative words) = 4 - 4 = **0**

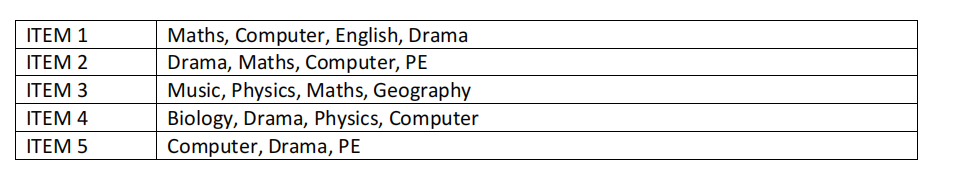
1. **Sentence 2:**

* "Liberating and transformative education can be inspiring and motivating, while alienating education can be unmotivating."
* **Positive words:** Liberating, transformative, inspiring, motivating (4 words)
* **Negative words:** Alienating, unmotivating (2 words)
* **Polarity of Sentence** = (Number of positive words) - (Number of negative words) = 4 - 2 = **2**

1. **Sentence 3:**

* "Enlightening education can be frustrating and demeaning, but it is still worthwhile in the long run."
* **Positive words:** Enlightening (1 word)
* **Negative words**: Frustrating, Demeaning (2 words)
* **Polarity of Sentence** = (Number of positive words) - (Number of negative words) = 1 - 2 = **-1**

## **Question 3**



**1) Maths, Computer => Drama**

Support = Frequency(Maths, Computer, Drama) / Total items

Frequency(Maths, Computer, Drama) = 0

**support** = 2/5 = 0.4

Confidence = Frequency(Maths, Computer, Drama) / Frequency(Maths, Computer)

**Confidence** = 2/2 = 1

Lift = Support / (Support(Maths) \* Support(Computer))

**Lift** = 0.4/0.6\*0.8 = 0.83

**2) Drama => Maths, Computer**

Support = Frequency(Drama, Maths, Computer) / Total items

**Support** = 2/ 5 = 0.4

Confidence = Frequency(Drama, Maths, Computer) / Frequency(Drama)

**Confidence** = 2 / 4 = 0.5

Lift = Support / (Support(Maths) \* Support(Computer))

Support(Maths) = 3/5 = 0.6

Support(Computer) = 3/5 = 0.6

**Lift** = 0.4 / (0.6\* 0.6) ≈ 1.111

**3) Computer, Drama => Physics**

Support = Frequency(Computer, Drama, Physics) / Total items

Frequency(Computer, Drama, Physics) = 1

**Support** = 1 / 5 = 0.2

Confidence = Frequency(Computer, Drama, Physics) / Frequency(Computer, Drama)

Frequency(Computer, Drama) = 4 (present in ITEM 2, ITEM 5)

Frequency(Computer, Drama, Physics) = 1 (present in ITEM 4)

**Confidence** = 1 / 4 = 0.25

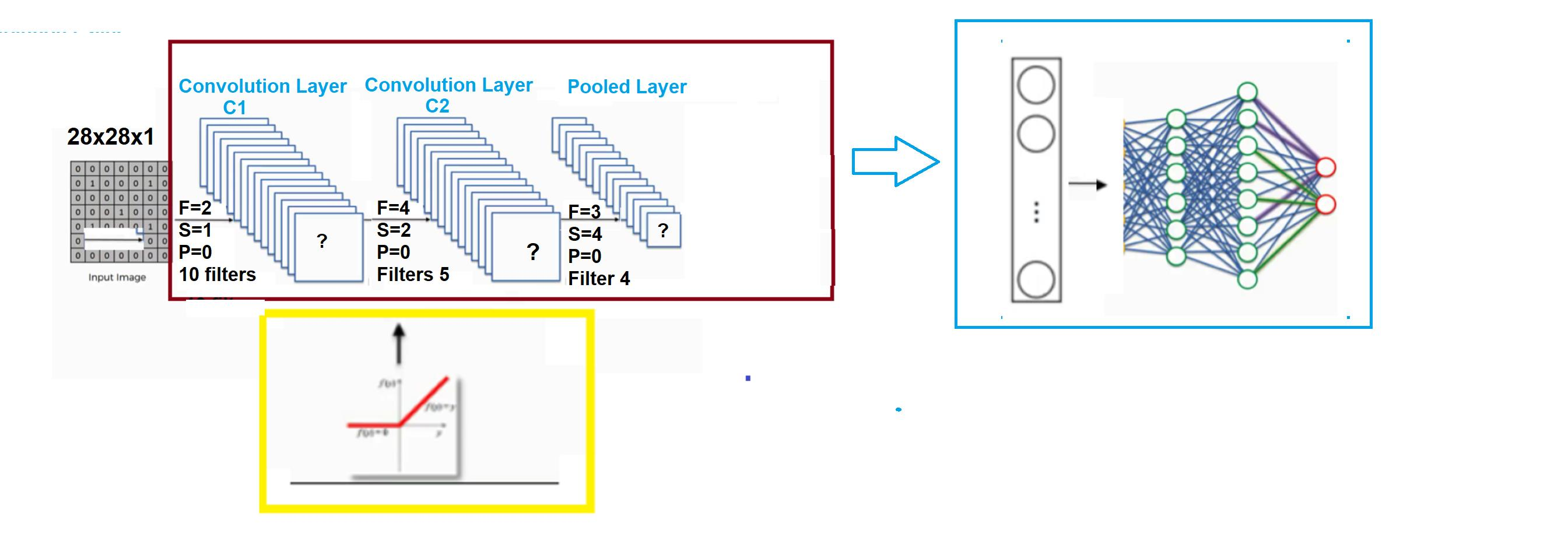
Lift = Support / (Support(Computer) \* Support(Drama))

Support(Computer) = 4 / 5 = 0.8

Support(Drama) = 4 / 5 = 0.8

Lift = 0.2 / (0.8 \* 0.8) ≈ 0.313

## **Question 4**



1. **In above figure, there are two units represented as red and blue . What**

**is name and functions of these red and blue unit in the above figure, and**

**what this network is called?**

**Answer:**

Red part -> Feature Extraction (consists of Convolution and pooling layers)

Blue Part -> Classification (consists of flatten and dense or fully connected layers)

This network is called **convolutional neural network** and is mostly used for image data.

1. **How many classes the above network can classify.**

**Answer:** 2 classes, as there are two neurons in the output layer.

1. **What is name of the function represent in yellow unit and what is its function in the network?**

**Answer:** This is **Relu** activation function and is used in hidden layers of network to introduced non linearity.

1. **Looking at red rectangle , where F represents filter size, S represents stride , P represents passing What would dimension of Layer C1 and C2 using the formula (n-f+2p)/s+1 (n-f+2p)/s+1.**

**Answer:** for layer C1  
=(28-2+2(0))/1+1=13 , so dimension of C1 will be 13\*13

For layer C2:

=(13-4+2(0))/2+1=3 , so dimension of C2 will be 3\*3

1. **Suppose you have 8000 training images to train this network , how many iteration it takes to complete 1 epoch of you use batch size of 500.**

**Answer:** 8000/500 = 16, it will take 16 iterations to complete one epoch.

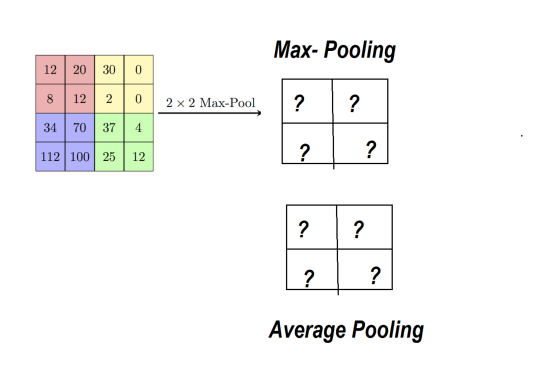
1. **Suppose you use Keras Library and sequential model to code this model. Following is the line of the code to create first two layers. What should you write for input\_shape.**

**model=Sequential()**

**model.add(Conv2D(32, kernel\_size = 3, activation='relu', input\_shape = (28\*28\*1) kernel\_initializer='he\_normal'))**

**model.add(Conv2D(32, kernel\_size = 3, activation='relu',kernel\_initializer='he\_normal'))**

1. **Look at the figure below and Calculate the max and average pooling.**



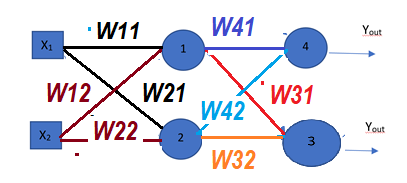
**Answer:**

**Max-pooling Average Pooling**

|  |  |
| --- | --- |
| **13** | **8** |
| **79** | **19.5** |

|  |  |
| --- | --- |
| **20** | **30** |
| **112** | **37** |

1. **.Look at the figure below**



Assuming

W11=1, W21=-1, W22=1, W12=-2,

W31=-2,W32=1, W41=2,W42=-3

Assume X1 =1 and X2=-1

Each of Node 1,2,3,4, uses following activation function

• f(v)=0 when v≥0;

• f(v)=1 otherwise

Calculate the output of Neuron1 ,Neuron2,Neuron3 and Neuron 4

**Answer:**

Neuron1 = V1= X1W11+X2W21 = 1\*1+(-1)\*(-1)=2 -> f(2)=0

Neuron2 = V2= X1W12+X2W22 = 1\*(-2)+(-1)\*(1)=-3 -> f(-3)=1

Neuron3 = V3= V1W31+V2W32 = 0\*(-2)+(1)\*(1)=1 -> f(2)=0

Neuron4 = V4= V1W41+V2W42 = 0\*2+(1)\*(-3)=-3 -> f(-3)=1

1. **Suppose you want to predict the Car price based on milage area (M) and Number of Years used (Y) and Engine Size €, suggest a Linear Regression formula to predict the car price (CP).**

*CP*=*β*0​+*β*1​×*M*+*β*2​×*Y*+*β*3​×*E*

*While* ***β*0** *is intercept and* ***β*1​, *β*2​, *β*3** *are the coefficients of variables M, Y and E respectively.*

1. **Differentiate between machine learning and Deep learning , how does the number of training samples effects the accuracy of classifiers in both of these cases.**

**Machine Learning:** Machine Learning involves using statistical algorithms and rules to learn patterns and make predictions from data. Examples of ML algorithms include, decision trees, naive bayes, svm etc. The ML algorithms can perform well on normal amount of data. So, we can achieve better accuracy while having sufficient no.of training examples.

**Deep Learning:** Deep Learning is sub-field of ML which uses artificial neural networks with a lot of layers to learn from a lot of data and then make predictions on unseen data. A large amount of data is required for deep learning. So, the number of training examples should be high for better generalization and accuracy.

1. **How do you measure the accuracy of KMean Clustering.**

**Answer:**

* **Inertia or Within-Cluster Sum of Squares (WCSS):** It measures the compactness of clusters. Lower inertia indicates better-defined clusters, where data points within clusters are closer to each other.
* **Visualization:** We can see the performance of clustering algorithms like Kmean clustering using scatter plot visualization.

1. **What is padding used for in CNN.**

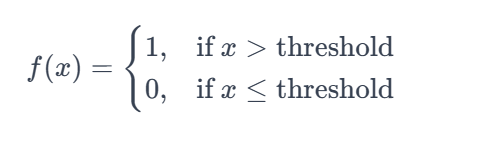
**Answer:** Padding in CNN is used for preserving the spatial information through the convolutional layers by adding extra pixels around the input data usually zeros. There are two options for padding

* **Same:** Same means zero padding which results in keeping original size of input in output and preserving information loss.
* **Valid:** Valid means no padding mean no extra pixels which results in smaller output than input

1. **What is the Stride in CNN.**

Stride is the step taken by the filter while moving on image features matrix. For example if stride is 1, it will skip only 1 row and 1 column while moving right and down respectively. and if stride is 2 then it will skip 2 rows and 2 columns while moving right and down respectively.

1. **Write down the activation function for perceptron.**

Step Function is used for perceptron.  


1. **Suppose we receive 2x2 matrix at the end of the CNN network , what would be the size of this matrix after flattening process before the fully connected layer.**

**Answer:** It will be 2\*2 = 4. We will get column vector having 4 values after flattening 2\*2 matrix.

1. **Differentiate between supervised , unsupervised and Reinforcement Learning.**

**Answer:** All these are the type of Machine Learning.

* **Supervised Learning:** Features and labels are given in dataset.
* **Unsupervised Learning:** Only features are given. No labels are defined.
* **Reinforcement Learning:** Learns on the basis of penalty and reward for the action taken.

1. **Write down steps required for the classification.**
2. Data Collection
3. Data Preprocessing
4. Feature Engineering
5. Splitting features and labels into training and testing sets
6. Model Selection
7. Training model on the preprocessed data
8. Evaluating model on test set
9. Saving model and making predictions on unseen data