Contents

[Question#01: 3](#_Toc494038971)

[Which to prefer? 3](#_Toc494038972)

[Problems with models A: 3](#_Toc494038973)

[Problems with models B: 3](#_Toc494038974)

[Model C with low Bias 3](#_Toc494038975)

[Question#02 3](#_Toc494038976)

[Algorithm and Reason: 4](#_Toc494038977)

[Question#03 4](#_Toc494038978)

[Part A) 4](#_Toc494038979)

[Part B) 4](#_Toc494038980)

[Question#04 4](#_Toc494038981)

[Part A) 4](#_Toc494038982)

[Part B) 5](#_Toc494038983)

[Question#05 5](#_Toc494038984)

[Question#06 5](#_Toc494038985)

[2 points in 2-D space 5](#_Toc494038986)

[3 points in 2-D space 5](#_Toc494038987)

[4 points in 2-D space 5](#_Toc494038988)

[Question#07 5](#_Toc494038989)

[PART A) 5](#_Toc494038990)

[PART B) 6](#_Toc494038991)

[Question#08 7](#_Toc494038992)

[Confusion Matrix: 7](#_Toc494038993)

[Type Error I (False positive) 7](#_Toc494038994)

[Type II Error (False negative) 7](#_Toc494038995)

[Question#09 7](#_Toc494038996)

[Question#10 7](#_Toc494038997)

[Induction Vs Deduction 7](#_Toc494038998)

# **Question#01:**

### **Which to prefer?**

Model A is better than Model B. There is only 5% difference between validation and testing accuracy which is reasonable. We may have some records in our test data with irregularities which penalty 5% low accuracy. Model B has low training accuracy and based on training testing accuracy is calculating so testing accuracy is reduced to half roughly in true sense.

### **Problems with models A:**

Model has sufficient data available for training. There may be missing values and outliers in the dataset. There may be problem with feature normalization. There can be feature selection problem. We may have selected features which are not relevant and effect our accuracy. This model shows high error on both the training and test sets, so the algorithm is suffering from high bias.

#### **Solution:**

* Increase the model complexity.
* Feature Transformation.
* Handle missing values and outliers.
* Feature selection and tuning.

### **Problems with models B:**

Model B has low training accuracy while higher test accuracy. It does mean model is underfitting. The large gap between training and test set errors, so the model is suffering from **high variance.**

#### **Solution:**

* More data will probably help
* Otherwise reduce the model complexity

### **Model C with low Bias**

Models with low bias are usually accurate on average but inconsistent. Models with low bias are usually complex models. Complex models lead to overfit. So low bias model can give us high training accuracy but low testing accuracy.

The solution with low bias model is to make our model less complex which will be less overfit. Decision tress are low bias and high variance algorithm. So, by apply pruning we can make decision trees less complex.

# **Question#02**

Yes, machine learning can be applied to this problem statement. We have a task, experience and performance measure metric.

* **Task** - Predict ways to deliver on time in hand.
* **Experience** – Previous Examples or data of rider’s rides.
* **Performance** – performance measure is if riders are delivering on time or not?

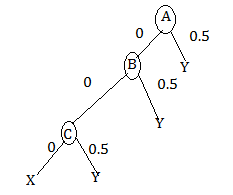
### **Algorithm and Reason:**

Decision trees can be utilized to handle this problem. An attribute can be chosen with minimum entropy. Once root node is calculated based on less entropy than entropy will be calculated for other nodes as well and select the one with minimum entropy and so on. We have different ways or tracks with traffic and conditions and we can choose one with minimum traffic and better condition. There can be there factors which effect the time to deliver. We need to consider all and build the decision tree. There can be so many ways, so decision tree can generalize the existing ways and predict the best one on run time.

# **Question#03**

### **Part A)**

Yes, such a tree is possible. Both classes are complement so such tree is possible in this specific case. Given below is an example which can handle case 1 and case 2. Below example provides a logical reasoning.



### **Part B)**

One can try decision tree with different attributes. We will use entropy to reduce the no of attributes. If model select the attribute with lowest entropy than attribute list can be reduced. The decision trees which give us the most information and those attributes which were not part of decision can be removed from the features. One can optimize this to select the best fit attribute which is most homogeneous with the data.

# **Question#04**

### **Part A)**

It is a supervised machine learning problem. Classification algorithm will be used. We have discrete output based on previous knowledge, so it’s a classification problem and not a regression problem.

### **Part B)**

It is a regression problem. that we are predicting the future income and we have labeled data for each year with continuous values. Data with continuous output values is best handled with regression algorithms.

# **Question#05**

No, Machine learning cannot be applied here. We do not have any inputs for this problem. For any machine learning problem, we need to have some inputs based upon model can predict output.

# **Question#06**

## 2 points in 2-D space

1. It can a circle equation and circle can shatter 2 points in 2-D plane.
2. It is equation for ellipse. We can shatter 2 points easily with ellipse.
3. It can be shatter by given hypothesis.

## 3 points in 2-D space

1. It is circle equation and circle can shatter 3 points in the 2-D plain.
2. It is ellipse equation and ellipse can shatter 3 points in the 2-D plain easily.
3. 3 points can be shatter by given hypothesis.

## 4 points in 2-D space

1. Circle can not shatter 4 points in 2-D space. There is a case [1010] with alternative positive and negative which is not possible with circle.
2. Ellipse can shatter 4 points in 2-D plain even when they are in alternative order.
3. 4 points can be not shatter by given hypothesis.

B model is more overfit. We have low bias in case of ellipses which leads to more overfit model.

# **Question#07**

### **PART A)**

#### **Automation Problems**

Automation is the linking of disparate systems and software in such a way that they become self-acting or self-regulating. In automation problems, we know the expected output already. We know if we provide these inputs, we will get this specific output.

##### **Automation Example:**

Solving a complex scientific formula is the best example of automation. One can implement a formula and some actions and provide inputs to get a specific output.

#### **Machine Learning Problems**

In machine learning problems, we do not know the expected output before the experiment. We need to perform an experiment to know the expected output. Machine learning problems usually have a task, previous data(experiments) and a performance measure.

##### **Machine learning Example**

If we have data for cancer for past 3 years. A model can be trained to predict the future cases of cancer. If a patient has cancer or not. It involves previous data and a performance measure.

### **PART B)**

In machine learning, the bias–variance tradeoff is the problem of simultaneously minimizing two sources of error that prevent supervised learning algorithms from generalizing beyond their training set. High bias can cause an algorithm to miss the relevant relations between features and target outputs (underfitting). High variance can cause an algorithm to model the random noise in the training data, rather than the intended outputs (overfitting).

# **Question#08**

## Confusion Matrix:

Confusion matrix, also known as an error matrix, is a specific table layout that allows visualization of the performance of an algorithm, typically supervised learning. Each row of the matrix represents the instances in a predicted class while each column represents the instances in an actual class.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Actual class** | | | | |
| **Predicted class** |  | **YES** | **NO** | **TOTAL** |
| **YES** | 80 | 20 | 100 |
| **NO** | 20 | 80 | 100 |
| **TOTAL** | 100 | 100 | 200 |

### **Type Error I (**[**False positive**](https://en.wikipedia.org/wiki/False_positive)**)**

Type one error is the false positive which is one positive actual but predicted positive by the model. It must be minimum for a better model predictor.

### **Type II Error (**[**False negative**](https://en.wikipedia.org/wiki/False_negative)**)**

Type 2 Error is false negative. All those records which were positive in actual but predicted negative by model. For our model to be better must have minimum II.

# **Question#09**

Attribute 3 and 4 provide us the maximum information. Petal length and petal width are the two attributes which are less overlap and hence provide us with the most information. we can delete attribute 1 and attribute 2 which are overlapped most and provide minimum information. Sepal length and sepal width will be deleted.

# **Question#10**

### Induction Vs Deduction

**Induction:**

Induction is the area with the larger number of methods. The goal in induction is to discover general concepts from a limited set of examples (experience). It is based on the search of similar characteristics or patterns among examples. It obtains general knowledge from specific information. The knowledge obtained is new. It’s not truth preserving i.e. It may invalidate the knowledge obtained.

**Example**:

The cost of goods was $1.00. The cost of labor to manufacture the time was $.50. The sales price of the item was $5.00; so, the item always provides a good profit.

**Deduction:**

Deductive reasoning obtains knowledge using well established methods or logic. The knowledge is not new and it is implicit in the initial knowledge. New knowledge cannot invalidate the knowledge already obtained. Its basis is mathematical logic. This is how mathematicians prove theorems from axioms. Proving a theorem is nothing but combining a small set of axioms with certain rules. Of course, this does not mean proving a theorem is a simple task, but it could theoretically be automated.

**Example:**

 We have these facts:

1. I watch a movie every Saturday
2. Today is Saturday

Deductive reasoning allows you to establish the *true statement* that I am going to watch today.