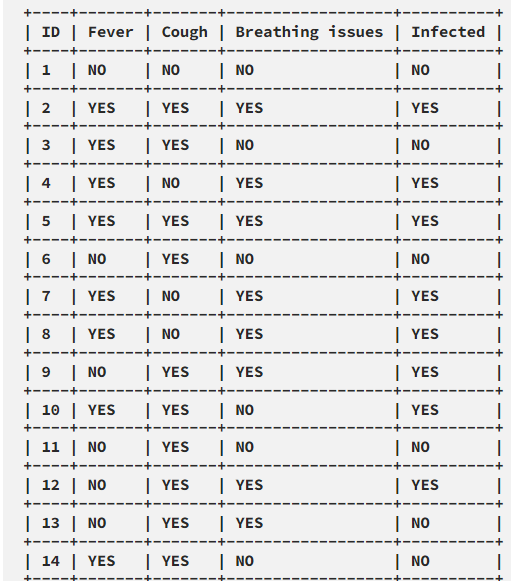
**Artificial intelligence Midterm exam TIME :100 M**

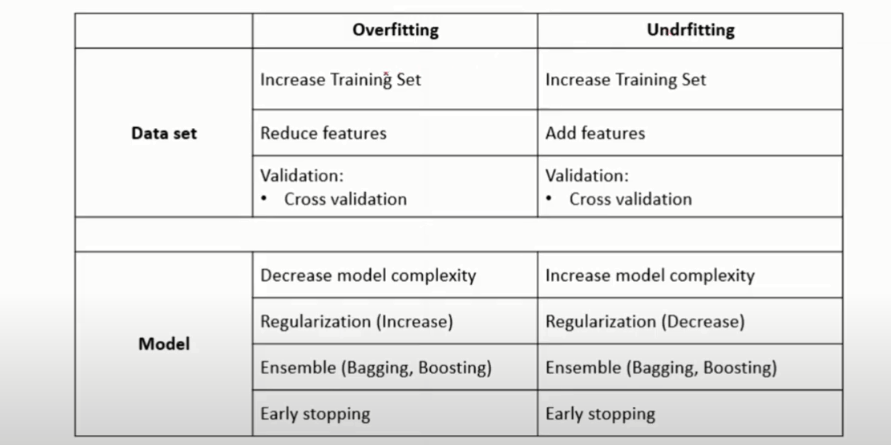
**Q1 [10]**

**A dataset of COVID-19 infection is given; create a Decision Tree using ID3 algorithm**



**Q2[12]**

1. **Suggest different Solution for underfitting – Overfitting.**

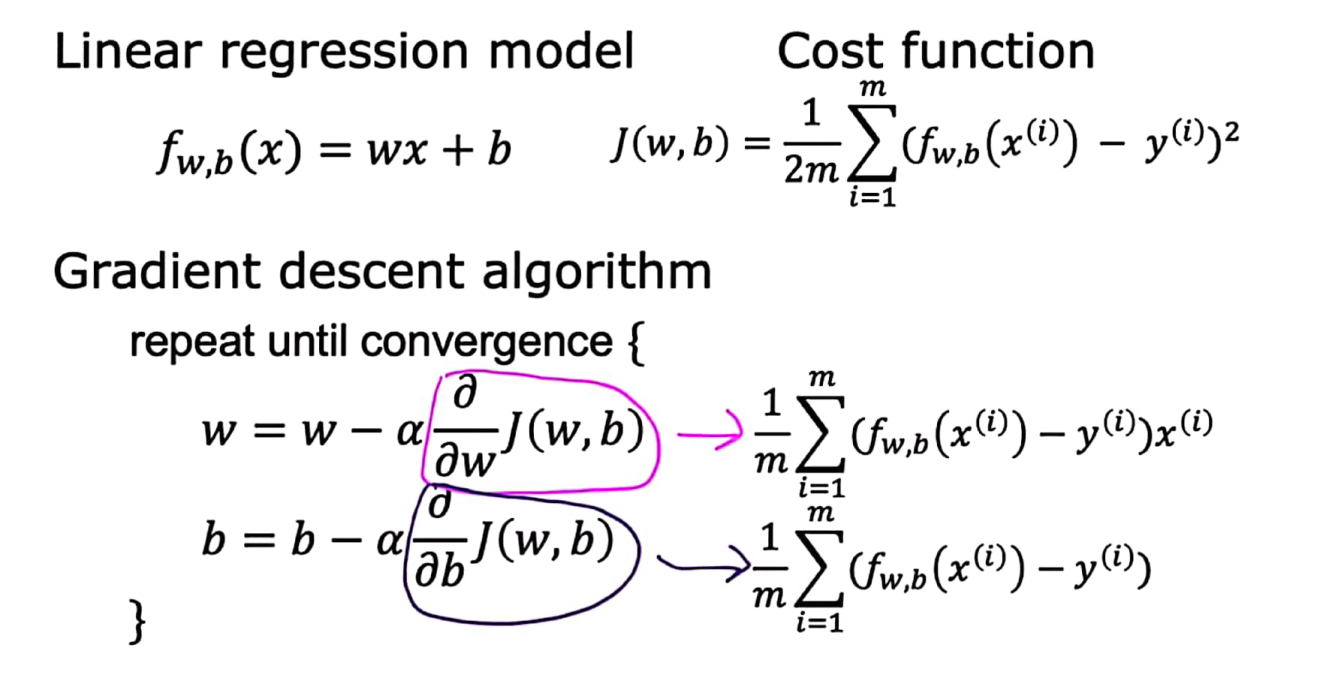
****

**b) Give two reasons for feature scaling.**

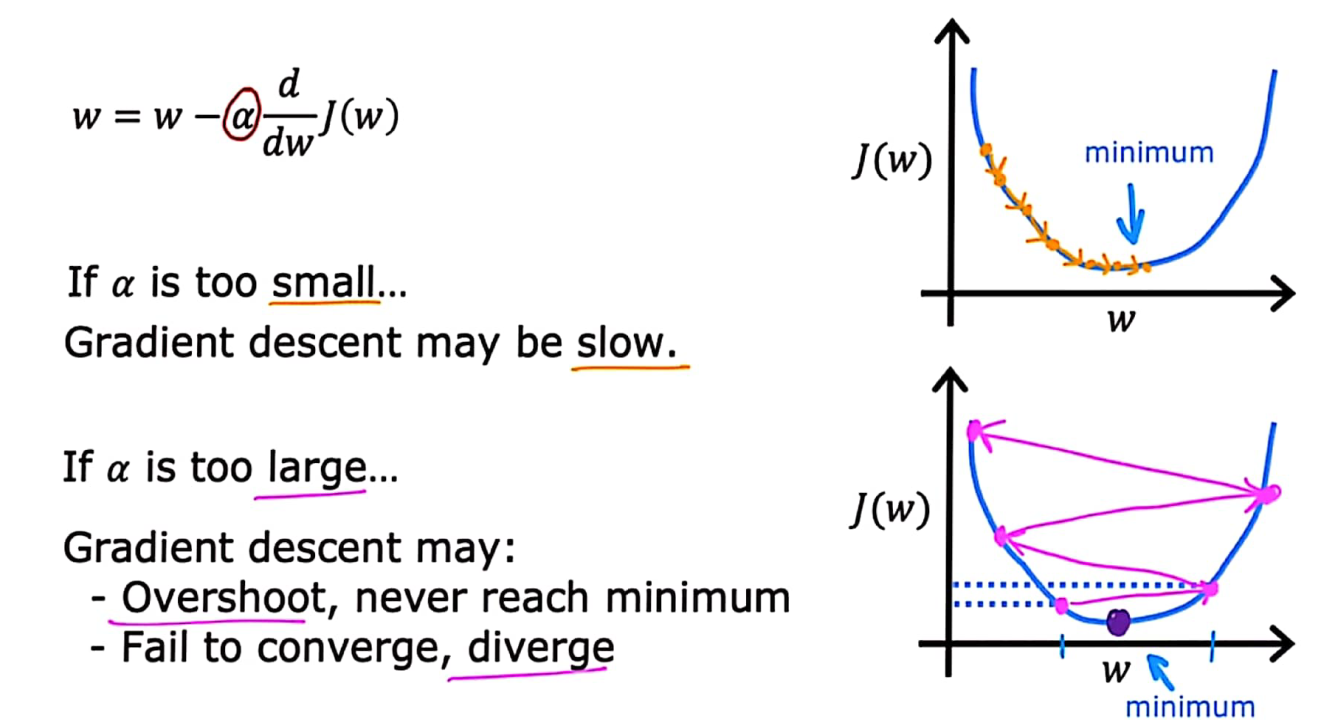
**1-Different scales , the cost surfaces is ellipse not circular , overshooting may be found, the convergence time is long.**

**2.IF the model is biased to size feature. The classification is not correct**

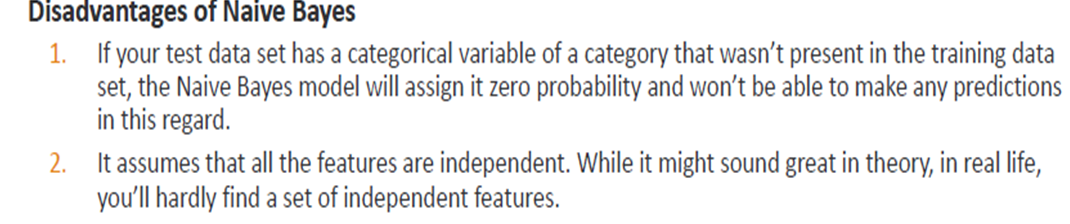
**c) Drive the equation of parameter update using gradient descent in linear regression.**

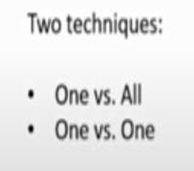
****

**D) Discuss the effect of choosing the learning rate value on model training process.**

****

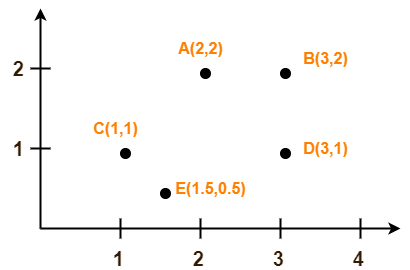
**e) Give two cases that naïve bayes algorithm isn't suitable for classification.**

**f) Can the model of logistic regression be used for multiclass classification? If yes show how.**

****

**Q3[8]**

**Use K-Means Algorithm to create two clusters - Assume A(2, 2) and C(1, 1) are centers of the two clusters.do one iteration only.**



**Q4 [10] For example, if for 100 observations, a model hits 40 (true positives), correctly rejects 43 (true negatives), raises 6 false alarms (false positives), and misses 11 (false negatives)**

1. ***Construct the confusion-matrix***
2. ***Find the accuracy***

## *****Find Recall*****

## *****Find Precision*****

## *****Find F1 Score*****

## Q4[10] choose the correct answer

1. **According to the American Society of Automotive Engineers (SAE) The automatic driving is divided**

**into Which level?**

**A. L1~L4 B. L1~L5 C. L0~L4 D. L0~L5**

**2- What is the performance of artificial intelligence in the stage of perceptual intelligence?**

**A. Machines begin to understand, think and make decisions like humans**

**B. Machines begin to calculate and transmit information just like humans**

**C. The machine starts to understand and understand, make judgments, and take some simple action**

**3- If a model has a large deviation on the test set and a small variance, it means that the model?**

**A. Overfitting B. May be overfitting may be underfitting**

**C. Just fit D. Underfitting**

**4- With a lot of sales data but no labels, companies want to identify VIP Customer, the following model Suitable?**

**A. Logistic regression B. SVM C. K-Means D. Hierarchical clustering**

**5- What are the common types of dirty data?**

**A. Malformed value B. Duplicate value C. Logically wrong value D. Missing value**

**6- What is the most important difference between batch gradient descent, mini-batch gradient descent, and stochastic gradient descent?**

**A. Gradient size B. Gradient direction C. Learning rate D. Number of samples used**

**7- The following about KNN Algorithm k, The value description is correct?**

**A. K The larger the value, the easier the model is to overfit.**

**B. K The larger the value, the smoother the segmentation surface of the classification**

**C. K Value is a hyperparameter D. can k Value is set to 0**

**8- In random forest, what strategy is used to determine the outcome of the final ensemble model?**

**A. Cumulative system B. Find the average C. Voting system D. Cumulative system**

**9- In supervised learning," Those who are near Zhu are red and those who are near Mo are black"Is used to describe which of the following models?**

**A. K-Means B. SVM C. KNN D. Neural Networks**

**10- The testing error will continue to decrease as the model complexity increases.**

**A. FALSE B. True**

Which machine learning models are trained to make a

series of decisions based on the rewards and feedback

they receive for their actions?

•A. Supervised learning

•B.Unsupervised learning

•C.Reinforcement learning

•D.All of the abov

Which machine learning models are trained to make a

series of decisions based on the rewards and feedback

they receive for their actions?

•A. Supervised learning

•B.Unsupervised learning

•C.Reinforcement learning

•D.All of the abov

**END OF EXAM**

## ****Accuracy****

Accuracy is the ratio of correct predictions to the number of predictions. It is given by:

https://www.academicianhelp.com/ckfinder/userfiles/files/image-20210615104709-1.png

For our example, the accuracy would be:

https://www.academicianhelp.com/ckfinder/userfiles/files/image-20210615104709-2.png

## ****Recall****

Recall is the ratio of true positive to actual positive. It gives the fraction of the actual positives that were hit. Since it measures how sensitive a model is to the positive class, it is also called sensitivity or True Positive Rate (TPR). It is given by:

https://www.academicianhelp.com/ckfinder/userfiles/files/image-20210615104709-3.png

For our example, the recall would be:

https://www.academicianhelp.com/ckfinder/userfiles/files/image-20210615104709-4.png

## ****Specificity****

This is the also called the True Negative Rate (TNR). It is the ratio of true negative to actual negative. It measures how correct the model is in not flagging the negative class. It is given by:

https://www.academicianhelp.com/ckfinder/userfiles/files/image-20210615104709-5.png

For our example, the specificity would be:

https://www.academicianhelp.com/ckfinder/userfiles/files/image-20210615104709-6.png

## ****Precision****

Precision measures how precise a model is when it classifies an observation as being positive. It is the ratio of true positive to predicted positive. It is also called Positive Predictive Value (PPV) and it is given by:

https://www.academicianhelp.com/ckfinder/userfiles/files/image-20210615104709-7.png

For our example, the recall would be:

https://www.academicianhelp.com/ckfinder/userfiles/files/image-20210615104709-8.png

## ****Negative Predictive Value****

This is long for NPV and it measures how accurate a model is in its negative predictions. It is the ratio of true negative to predicted negative. It is given by:

https://www.academicianhelp.com/ckfinder/userfiles/files/image-20210615104709-9.png

For our example, the recall would be:

https://www.academicianhelp.com/ckfinder/userfiles/files/image-20210615104709-10.png

## ****F1 Score****

Accuracy may be misleading when the class is heavily unbalanced because its value may be inflated by the majority class. F1 score is a robust metric to heavily unbalanced data because it penalizes misclassification more. It is the harmonic mean of precision and recall and it is given by:

https://www.academicianhelp.com/ckfinder/userfiles/files/image-20210615104709-11.png

For our example, the recall would be:

https://www.academicianhelp.com/ckfinder/userfiles/files/image-20210615104709-12.png

**Solution-**

We follow the above discussed K-Means Clustering Algorithm.

Assume A(2, 2) and C(1, 1) are centers of the two clusters.

**Iteration-01:**

* We calculate the distance of each point from each of the center of the two clusters.
* The distance is calculated by using the euclidean distance formula.

The following illustration shows the calculation of distance between point A(2, 2) and each of the center of the two clusters-

**Calculating Distance Between A(2, 2) and C1(2, 2)-**

Ρ(A, C1)

= sqrt [ (x2 – x1)2 + (y2 – y1)2 ]

= sqrt [ (2 – 2)2 + (2 – 2)2 ]

= sqrt [ 0 + 0 ]

= 0

**Calculating Distance Between A(2, 2) and C2(1, 1)-**

Ρ(A, C2)

= sqrt [ (x2 – x1)2 + (y2 – y1)2 ]

= sqrt [ (1 – 2)2 + (1 – 2)2 ]

= sqrt [ 1 + 1 ]

= sqrt [ 2 ]

= 1.41

In the similar manner, we calculate the distance of other points from each of the center of the two clusters.

Next,

* We draw a table showing all the results.
* Using the table, we decide which point belongs to which cluster.
* The given point belongs to that cluster whose center is nearest to it.

|  |  |  |  |
| --- | --- | --- | --- |
| **Given Points** | **Distance from center (2, 2) of Cluster-01** | **Distance from center (1, 1) of Cluster-02** | **Point belongs to Cluster** |
| A(2, 2) | 0 | 1.41 | C1 |
| B(3, 2) | 1 | 2.24 | C1 |
| C(1, 1) | 1.41 | 0 | C2 |
| D(3, 1) | 1.41 | 2 | C1 |
| E(1.5, 0.5) | 1.58 | 0.71 | C2 |

From here, New clusters are-

**Cluster-01:**

First cluster contains points-

* A(2, 2)
* B(3, 2)
* E(1.5, 0.5)
* D(3, 1)

**Cluster-02:**

Second cluster contains points-

* C(1, 1)
* E(1.5, 0.5)

Now,

* We re-compute the new cluster clusters.
* The new cluster center is computed by taking mean of all the points contained in that cluster.

**For Cluster-01:**

Center of Cluster-01

= ((2 + 3 + 3)/3, (2 + 2 + 1)/3)

= (2.67, 1.67)

**For Cluster-02:**

Center of Cluster-02

= ((1 + 1.5)/2, (1 + 0.5)/2)

= (1.25, 0.75)

This is completion of Iteration-01.

Next, we go to iteration-02, iteration-03 and so on until the centers do not change anymore.

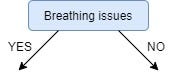
**https://towardsdatascience.com/decision-trees-for-classification-id3-algorithm-explained-89df76e72df1**

**Entropy(S) = — (8/14) \* log₂(8/14) — (6/14) \* log₂(6/14) = 0.99**

The block, below, demonstrates the calculation of Information Gain for **Fever.**

# total rows **|S| = 14For v = YES, |Sᵥ| = 8  
Entropy(Sᵥ) = - (6/8) \* log₂(6/8) - (2/8) \* log₂(2/8) = 0.81For v = NO, |Sᵥ| = 6  
Entropy(Sᵥ) = - (2/6) \* log₂(2/6) - (4/6) \* log₂(4/6) = 0.91**# Expanding the summation in the IG formula: **IG(S, Fever) = Entropy(S) - (|S**ʏᴇꜱ**| / |S|) \* Entropy(S**ʏᴇꜱ**) -   
(|S**ɴᴏ**| / |S|) \* Entropy(S**ɴᴏ**)∴** **IG(S, Fever) = 0.99 - (8/14) \* 0.81 - (6/14) \* 0.91 = 0.13**

**IG(S, Cough) = 0.04  
IG(S, BreathingIssues) = 0.40**



**IG(Sʙʏ, Fever) = 0.20  
IG(Sʙʏ, Cough) = 0.09** 