**MOCK INTERVIEW**

Name: M S Narendhar

Roll No: CB.EN.U4CSE19031

**Q1). How to find loop in linked list**

**Ans:** We find the loop using Floyd’s-cycle finding algorithm.In this method we have 2 pointers P and Q which pointed to head node.Then we move pointer P by one step and pointer Q by 2steps.If these 2 pointers meet at same node then there is loop otherwise there is not a loop.

Algorithm:

Def findloop(L):

P=L.head

Q=L.head

While(P && Q && Q.next)

{

P=P.next

Q=Q.next.next

If(P==Q)

{

Return

}

}

**Q2). Is it possible to construct a tree with pre-order traversal alone.**

**Ans:** To build a binary tree we need 2 traversals.eg pre and post or pre and in order.

This is because let’s take N elements. Therefore there are N! possibilities for the n elements where the possible no of trees are much larger. Thus from pigeon hole principle there is at least one traversal generates 2 trees. So we cannot reconstruct the tree from single traversal.

**Q3). How to find the time complexity for recursive algorithms**

**Ans:** First we need to identify the input size and smaller sub problems. Then we need to find recurrence relation for the algorithm. Then we need to solve the recurrence relation using Recursion tree method or using master’s theorem.

**Q4). Recurrence relation and time complexity for Binary Search algorithm?**

**Ans:** T(n) = T(n/2) + Ɵ(1)

Here a=1 and b=2,using master theorem, n^log a base b =n^0 =>1 and f(n)= Ɵ(1)

Therefore using case 2 of master theorem,

T(n) = Ɵ( n^log a base b \* log n ) => Ɵ(log n)

Therefore time complexity is Ɵ(log n).

**Q5). For the given dataset to predict the total what kind of model will preferred ?**

**Ans:** Since the data points are continuous, we will use regression model. Regression is a supervised machine learning technique which is used to predict continuous values. The ultimate goal of the regression algorithm is to plot a best-fit line or a curve between the data. The three main metrics that are used for evaluating the trained regression model are variance, bias and error. If the variance is high, it leads to overfitting and when the bias is high, it leads to underfitting.

**Q6).For the given dataset if we need to find the fast learners and slow learners then what kind of model will preferred ?**

**Ans:** Here we need to split it on only 2 groups, so here classification model is used. A classification model attempts to draw some conclusion from observed values. Given one or more inputs a classification model will try to predict the value of one or more outcomes. Outcomes are labels that can be applied to a dataset. For example, when filtering emails “spam” or “not spam” (also known as “ham”, < seriously, look it up if you don’t believe me), when looking at transaction data, “fraudulent”, or “authorize.

**Q7). How to remove outliers in the given dataset**

**Ans:** Outliers are those data points which differs significantly from other observations present in given dataset. It can occur because of variability in measurement and due to misinterpretation in filling data points.

Outlier can be of two types: Univariate and Multivariate. Above, we have discussed the example of univariate outlier. These outliers can be found when we look at distribution of a single variable. Multi-variate outliers are outliers in an n-dimensional space.

a) Z-score method:

z= datapoint -mean/standard deviation

Here, we normally define outliers as points whose modulus of z-score is greater than a threshold value. This threshold value is usually greater than 2 (3 is a common value).

b) I.Q.R method:

In this method by using Inter Quartile Range (IQR), we detect outliers. IQR tells us the variation in the data set. Any value, which is beyond the range of -1.5 x IQR to 1.5 x IQR treated as outliers.

* Q1 represents the 1st quartile/25th percentile of the data.
* Q2 represents the 2nd quartile/median/50th percentile of the data.
* Q3 represents the 3rd quartile/75th percentile of the data.
* (Q1–1.5IQR) represent the smallest value in the data set and (Q3+1.5IQR) represent the largest value in the data set