**+Western Union**

**Technical Round 1;**

Project:

1. Which data you have used and how you got the data?
2. **Were you involved in model deployment?**
3. **Do you have any idea about inference or what happens when we build a pickle file? What happens next?**
4. your problem statement is classification then why u used regression?
5. After using the regression algorithm, how were you classified into defective and good ICs?
6. **can we have different probabilities for classifying the wafer?**
7. if you said using probability, we can classify the wafer then why you haven't gone with classification?
8. What is your target variable, is it regression or classification?9] why you went ahead with linear regression and why?

Linear Regression?

1] What is a linear regression algorithm?

1] What are the assumptions of linear regression and how do you implement them practically?

1. what is VIF and explain it in plain English & what is the VIF formula?
2. Are there any other cost functions other than MSE?
3. What are the methods to find the best parameters(weights)?
4. When will you use OLS and Gradient descent?
5. What are the input parameters of gradient descent?
6. **Why is everyone using Gradient descent these days & why is there so much hype about gradient descent?**

Ans - When we have high dimensional data the gradient descent is fast.

1. what are the input parameters we pass to gradient descent if we are implementing from scratch?

Ans- **learning rate, weights, independent sample.**

1. Initially what do you do while building gradient descent from scratch?

Ans - Initially we assign the weights.

1. What is the stop condition for gradient descent?

Ans- We stop gradient descent when the current loss function is not changing or it is increasing.

1. What is the formula of gradient descent?
2. Gradient descent will minimize or maximize the cost function?

**Logistic Regression:**

1. Explain thoroughly logistic regression?
2. Can we use linear regression for classification, answers in Yes or NO, If yes then how and why?
3. What is the cost function of logistic regression & its formula & what is the name of the field of the formula?
4. What is a sigmoid function?
5. What is the probability formula in logistic regression?
6. Formula of cross entropy & what is the probability in cross entropy and what is the formula of probability in cross entropy? 7]

Python code:

Return the first element to repeat in list: input =[a,b,c,d,e,b,f,g] Output = b

input =[a,b,c,d,e,a,b,f,g] **Managerial Round:**

**Walk me through your resume?**

**Explain the VM project thoroughly?**

**You said you had 500 features, how did you handle them? Which data quality test you checked?**

**Now after preprocessing you said you had 36 features now from here also can we reduce them?**

**Ans - I told him that we can use PCA for feature**

**reduction then he told if we use PCA we would lose the variance of the features because PCA can't have all the variance of the original feature so do you have any other idea than the PCA?**

**Again I told him that we can take features for ex x1 and x2 and we can combine them then he was fine with it.**

**Explain linear regression and its assumption.**

**How do you check linear regression assumptions other than charts?**

**Answer: Not able to answer**

**How do you handle multicollinearity and what is r - square and VIF?**

**Do you have any idea about feature selection methods like forward and backward selection?**

**Answer - I told him I know them but haven't used them in the project then he was fine with it.**

**Do you have any idea about weighting algorithms and feature clustering? Answer - Told him don't have any idea.**

**What are underfitting and overfitting? How do you handle both of them when they occur?**

**What is data imbalance? How do you handle them?**

**For example, in financial fraud activities we don't have much data about fraud activities so what would do in this case While building a model?**

**Answer: In this, it doesn't make sense to do understanding and oversampling of the data because we are just replicating the sample to make the class equals but it is not going to add any value to the model.**

**So instead, we use the SMOTE method to generate synthesized data using knn by considering near neighbors.**

**Does overfitting or underfitting happen on the imbalance datasets?**

**Answer: I think it depends on the data if we see there are higher chances that the model will predict the over-sample labels so this is overfitting and the model is not able to learn to undersample labels so it is underfitting.**

**What is cross-validation? What is the ideal value for k?**

**Suppose you have two datasets with samples of 1000 and 100 now how many fold u will use on each dataset?**

**Suppose you have two features and they are highly correlated but you can't remove them because business people told you to add them. How would you handle this situation? Answer: we can combine these two features. He was impressed** 😂

**Why do you want to join our company?**

**Do you have any offers in hand?**

**Do you have any questions for me?**

**Why are you leaving your current job?**

**Python code:**

**Print the square of each element in case char prints it twice.**

**[1,5,9,7,a,d,cc,gg]**

**SQL code :**

**Write a SQL query to get the highest-selling car type.**

**Table1**

**Car name, Month, Quantity sold**

**Table 2:**

**Car name, Car Type**

**Suppose we have two datasets**

**Dassault Systemes**

1.diff list n tuple which is fast n why

2.n prime number code

3.

a=[]

b=[]

a.append(i)

b.append(a)

what is concept called

4.tp=(2,3,4,[5,6]) can we add element in list

5.what is static and upstrack method

6.encapsulation polymorphism

7.constuctor

8.class

9.methd

10.generator

11.lambda

12.closure

pandas:

1.agg()

2.applymap()

3.what is average salary of employee conider under 3 standard deviation data then take avg

4.join vs merge

round 2

1.sum of two number is = given target find out those number's index efficient way

2.given a string find the 2nd most occured char

3.puzzle

why do you want to join DS

if someone give you 15 LPA would you join