Daily Machine Learning Interview Questions

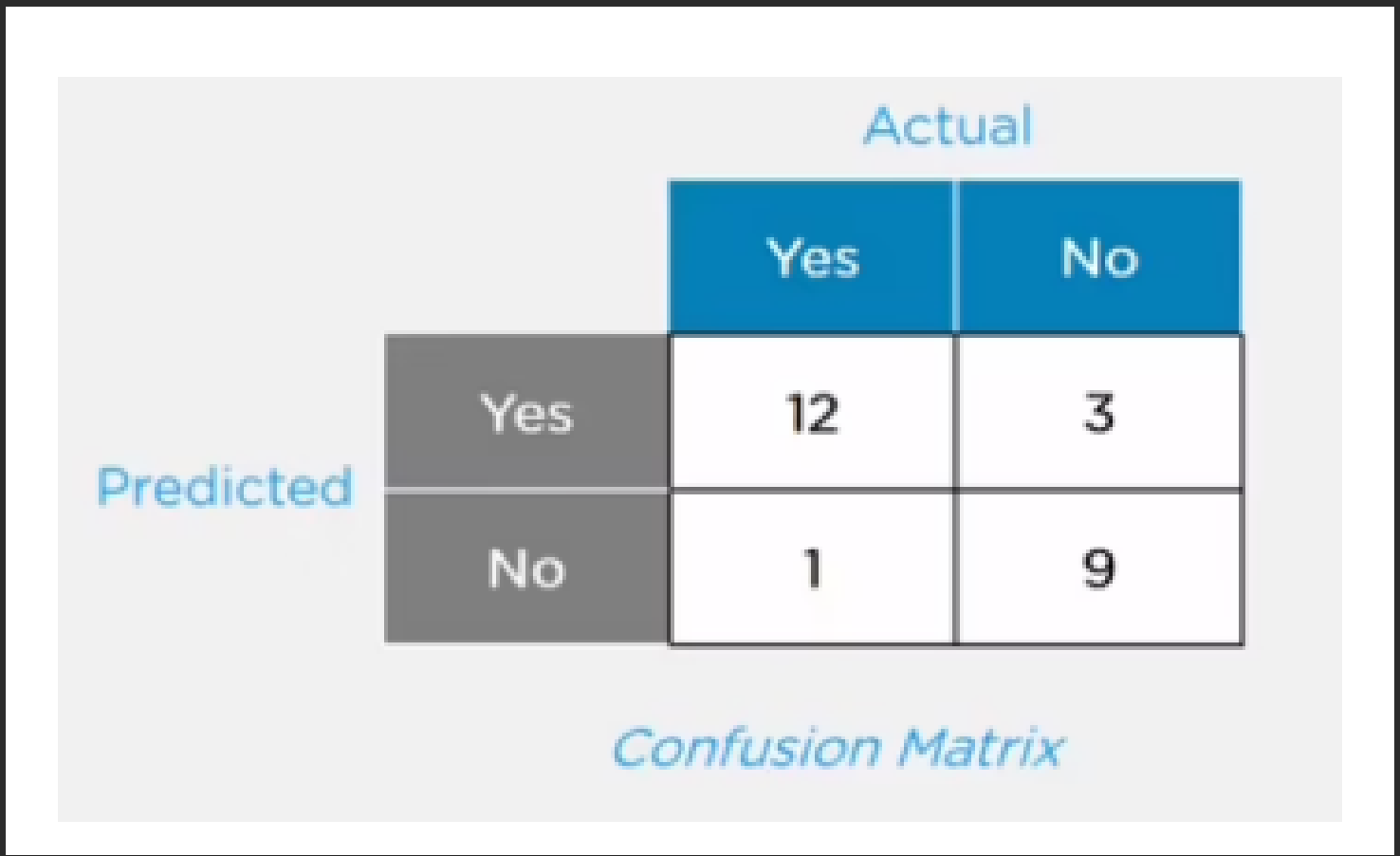
31. Explain the Confusion Matrix with Respect to Machine Learning Algorithms

A confusion matrix (or error matrix) is a specific table that is used to measure the performance of an algorithm. It is mostly used in supervised learning; in unsupervised learning, it’s called the matching matrix

The confusion matrix has two parameters:

- Actual
- Predicted

It also has identical sets of features in both of these dimensions



Here,
For actual values:

Total Yes = 12+1 = 13
Total No = 3+9 = 12

Similarly, for predicted values:
Total Yes = 12+3 = 15
Total No = 1+9 = 10

For a model to be accurate, the values across the diagonals should be high. The total sum of all the values in the matrix equals the total observations in the test data set.

For the above matrix, total observations = 12+3+1+9 = 25

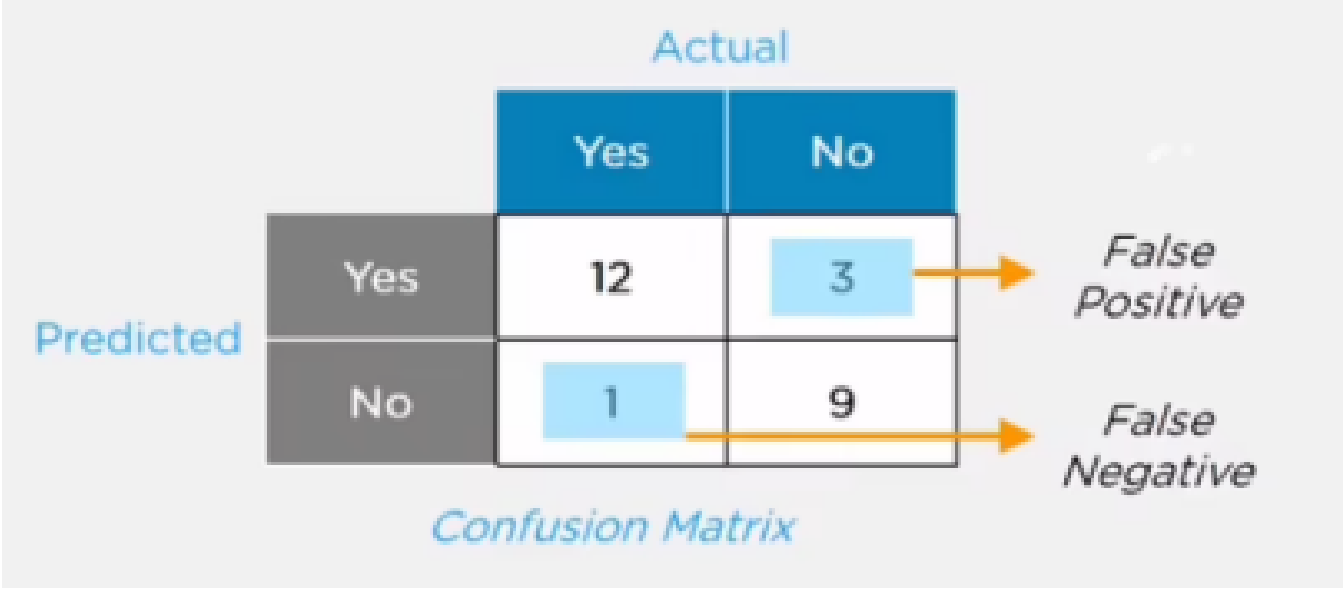
Now, accuracy = sum of the values across the diagonal/total dataset

= (12+9) / 25
= 21 / 25
= 84%

32. What Is a False Positive and False Negative and How Are They Significant?

False positives are those cases that wrongly get classified as True but are False. False negatives are those cases that wrongly get classified as False but are True.

In the term ‘False Positive,’ the word ‘Positive’ refers to the ‘Yes’ row of the predicted value in the confusion matrix. The complete term indicates that the system has predicted it as a positive, but the actual value is negative.



So, looking at the confusion matrix, we get:
False-positive = 3
True positive = 12

Similarly, in the term ‘False Negative,’ the word ‘Negative’
refers to the ‘No’ row of the predicted value in the confusion
matrix. And the complete term indicates that the system has
predicted it as negative, but the actual value is positive.

So, looking at the confusion matrix, we get:
False Negative = 1
True Negative = 9

33. Define Precision and Recall.

Precision :

Precision is the ratio of several events you can correctly recall to the total number of events you recall (mix of correct and wrong recalls). Precision = (True Positive) / (True Positive + False Positive)

Recall :

A recall is the ratio of the number of events you can recall the number of total events. Recall = (True Positive) / (True Positive + False Negative)

34. What do you understand by Type I vs Type II error?

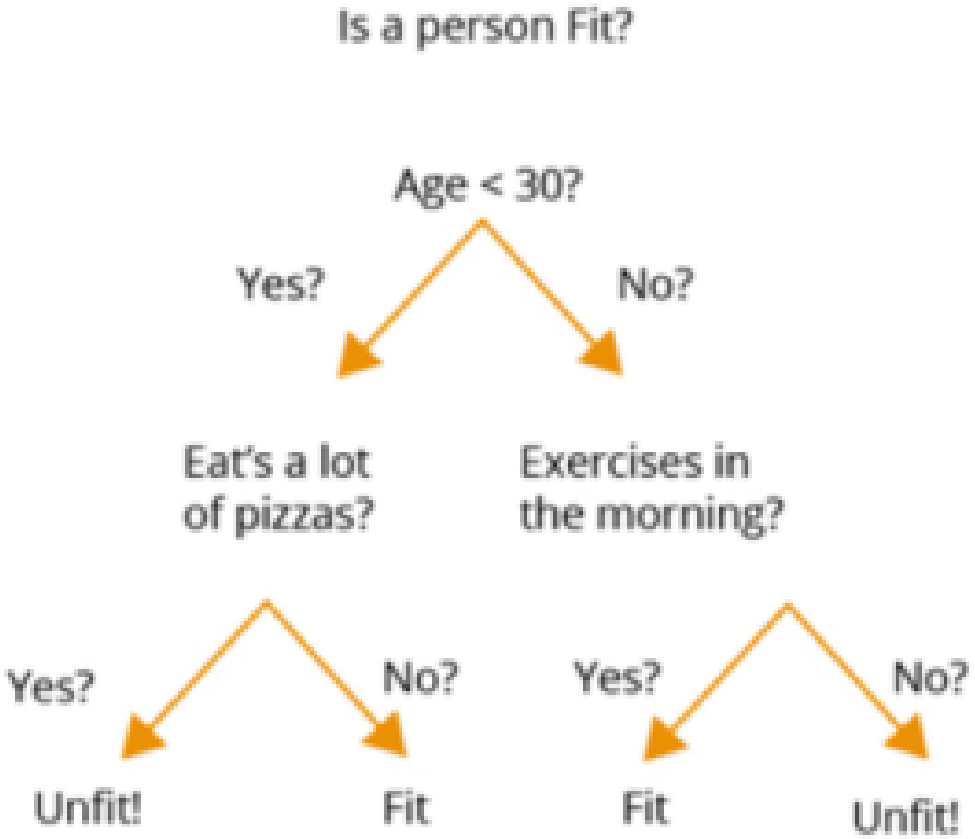
Type I Error: Type I error occurs when the null hypothesis is true and we reject it.

Type II Error: Type II error occurs when the null hypothesis is false and we accept it.

		reality	
		H_0 = True	H_0 = False
Conclusion	H_0 is not rejected	OK	Type II error
	H_0 is rejected	Type I error	OK

35. What is a Decision Tree in Machine Learning?

A decision tree is used to explain the sequence of actions that must be performed to get the desired output. It is a hierarchical diagram that shows the actions.


Is a person Fit?
Age < 30?
Yes?
No?
Eat's a lot of pizzas?
Exercises in the morning?
Yes?
No?
Yes?
No?
Unfit!
Fit
Fit
Unfit!

An algorithm can be created for a decision tree on the basis of the set hierarchy of actions.

In the above decision-tree diagram, a sequence of actions has been made for driving a vehicle with or without a license.

Thank You