

Confusion Matrix:

Will be given:

Number of predictions

Number of positive predictions

Number of negative predictions

Will be asked:

Calculate precision and recall of 100 predictions, 60 True predictions, 40 False predictions, 28 actually were true, 14 actually were negative.

What is True Positive and True Negative?

Will be asked how you would rate a prediction model in a fake organization with 90 percent accuracy, and 84 percent recall

Accuracy = $(TP + TN) / (TP + TN + FP + FN)$

Error = $FP + FN / TP + TN + FP + FN$

Precision = $TP / TP + FP$ (When model predicted true, how often was it right?)

Recall = $TP / TP + FN$ (when the class was actually true, how often did the classifier get it right?)

Statistics

-A p-value is a measure of the probability that an observed difference could have occurred just by random chance.

-The lower the p-value, the greater the statistical significance of the observed difference.

-P-value can be used as an alternative to or in addition to pre-selected confidence levels for hypothesis testing.

- if the P value is less than 0.05, you reject the null hypothesis.

Will be asked to demonstrate an example of when you would use a T test or a p-value and asked to explain as if you were talking to a non tech board member.

Model Validation Techniques:

Splitting your data. ...

Discuss your options, such as follows:

k-Fold Cross-Validation (k-Fold CV) ...

Leave-one-out Cross-Validation (LOOCV) ...

Nested Cross-Validation. ...

Time Series CV. ...

Comparing Models.

NN Regularization: Reduce Overfitted/Variance by making it less complex.

Lower fit on training = better ability to generalize on unseen data.

Drop out: The probability of keeping each node is set at random. You only decide of the threshold: a value that will determine if the node is kept or not.

L1: Lasso Regression (Least Absolute Shrinkage and Selection Operator) adds “absolute value of magnitude” of coefficient as penalty term to the loss function.

L2 - large weights are penalized (Ridge Regression) (most common)

Ridge regression adds “squared magnitude” of coefficient as penalty term to the loss function. Here the highlighted part represents L2 regularization element.

Be able to talk about which regularization you would prefer in a situation, and why.

General

Be able to talk about a dataset you care about/are passionate about. Provide a project you are doing with that dataset. Share insights.