1. what cross-validation is and when it is usually used.

Cross validation is a technique of evaluating a model by building several models on different parts of the data. It means if the if we split the data into 5 parts, So we will run 5 models, and every time 1 part of the 5 parts will be test set. And it is used when we want to get the range of possible accuracies of the data. Also we can find the overfitting of the model.

1. the differences of cross-validation strategies in evaluating your chosen machine learning model's performance.

I have built a model using Random Forest Algorithm. For standard k-fold cross-validation with default folds the accuracy is 0.48, for standard k-fold cross-validation with 10 splits its, 0.50, with stratified k-fold cross-validation it is again 0.501, for Leave-one-out cross-validation its and finally with Shuffle-split cross-validation it is 0.48.

1. the purpose for splitting the data set into training set, validation set, and test set for certain machine learning models

When we build a model using certain data, that model is running in our local environment. The model which is built on the already known data will not work in global environment. So purpose of splitting the data into into train, test, and validation to make the model work better in gloabal environment. So Once after creating a model is train data we can verify it using validation data, and validate using test data.

1. what is false positive and false negative in binary classification, and why accuracy alone is not a good measure for the machine learning algorithms.

False positive is basically like type I error, type I error is rejecting null hypothesis when is actually true. False negative is like type II error I basically fails to reject null hypothesis when it is actually false. False negative and False Positives are really important is case of cetain classification examples. Accuracy alone is not a good measure of how good the model is. Recall and precision are more important in classification models. Both the measure basically count in the weightage of all the classes in terms of prediction, so that we get clear picture how model exactly predicting all the classes.