Stirring Minds

MACHINE LEARNING TASK-1:

Q. You are assigned to a project where you are working on a datasets provided with Career guidance which contains various details of the students and their preferred career choices. How do you select the important variables? Explain your method in detail. (You are free to assume the variables as needed).

Ans. While work on a dataset we can select the important variables in the following ways.

Initially we take the whole data set and then apply step wise regression.

Step wise Regression includes (choose 1):

➔ Backward Elimination
➔ Forward Selection
➔ Bidirectional Elimination

Finally score comparison

I prefer using backward elimination usually since it is the fastest.

In backward Elimination, there are 5 steps

Step 1. Select a significance level

Step 2. Fit the full model with all possible predictors

Step 3. Consider the predictor with highest P-value. If P>Significant level, go to step 4 else step 5

Step 4. Remove the predictor

Step 5. Fit model without this variable

In this way we can select the important variables we are going to need for our model.