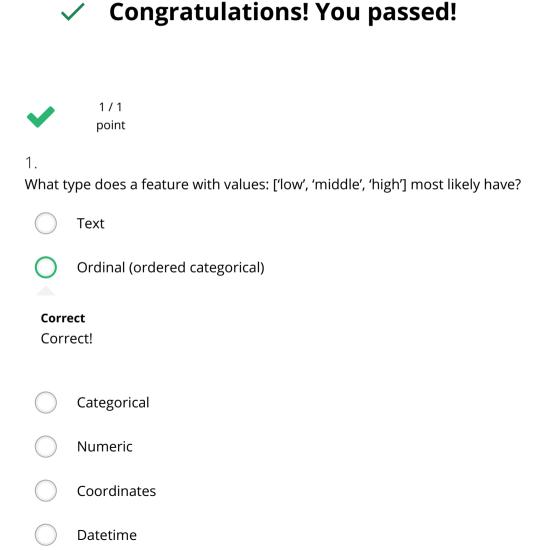


6/6 points (100%)

Next Item





Feature preprocessing and generation with respect to models

6/6 points (100%)



2/2 points

Practice Quiz, 4 questions

2

Suppose you have a dataset X, and a version of X where each feature has been standard scaled.

For which model types training or testing quality can be much different depending on the choice of the dataset?

Nearest neighbours

Correct

Correct! The reason for it is that the scale of features impacts the distance between samples. Thus, with different scaling of the features nearest neighbors for a selected object can be very different.



Neural network

Correct

Correct! There are two reasons for this: first, amount of regularization applied to a feature depends on the feature's scale. Second, optimization methods can perform differently depending on relative scale of features.



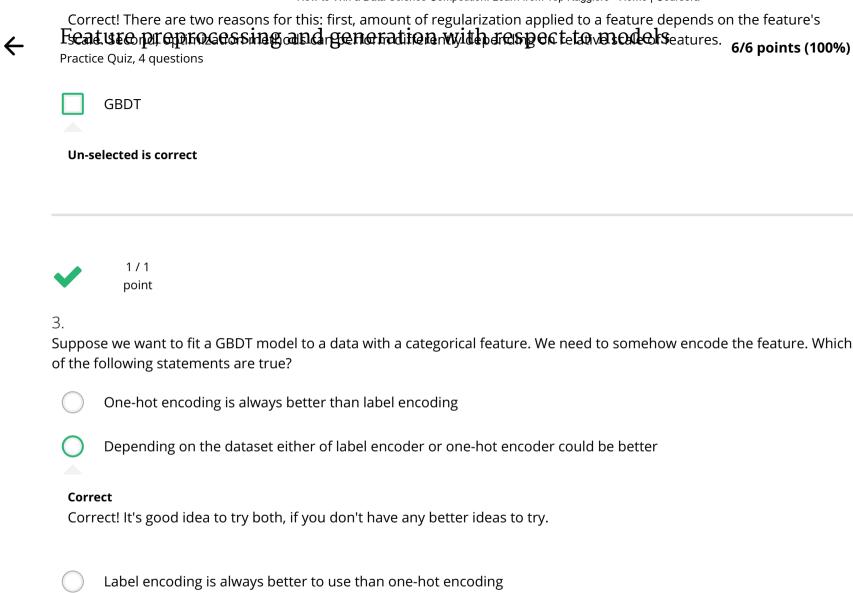
Random Forest





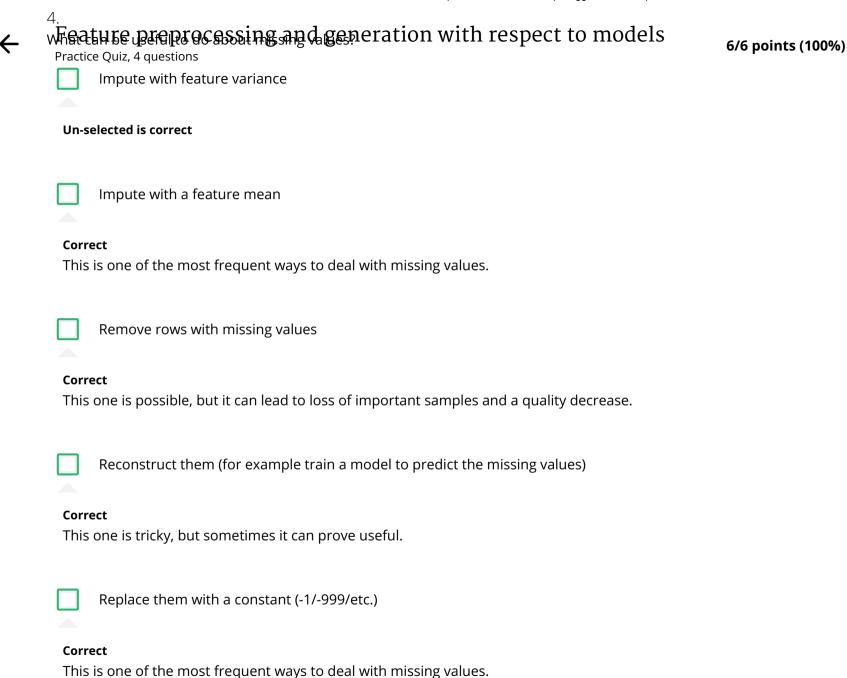
Linear models

Correct





2/2 points





Nothing, but use a model that can deal with them out of the box Feature preprocessing and generation with respect to models

6/6 points (100%)

Practice Quiz, 4 questions
Correct

Some models like XGBoost and CatBoost can deal with missing values out-of-box. These models have special methods to treat them and a model's quality can benefit from it.



Apply standard scaler

Un-selected is correct





