

1. (True/False) Skipping data points (i.e., skipping rows of the data) that have missing features only works when the learning algorithm we are using is decision tree learning. 1 point

☐ True
☒ False

2. What are potential downsides of skipping features with missing values (i.e., skipping columns of the data) to handle missing data? 1 point

- ☒ So many features are skipped that accuracy can degrade
☐ The learning algorithm will have to be modified
☐ You will have fewer data points (i.e., rows) in the dataset
☒ If an input at prediction time has a feature missing that was always present during training, this approach is not applicable.

3. (True/False) It's always better to remove missing data points (i.e., rows) as opposed to removing missing features (i.e., columns). 1 point

☐ True
☒ False

4. 1 point

Consider a dataset with N training points. After imputing missing values, the number of data points in the data set is

☐ $2 * N$
☒ N
☐ $5 * N$

5. 1 point

Consider a dataset with D features. After imputing missing values, the number of features in the data set is

☐ $2 * D$
☒ D
☐ $0.5 * D$

6.

1 point

Which of the following are always true when imputing missing data? Select all that apply.

- ☒ Imputed values can be used in any classification algorithm
- ☒ Imputed values can be used when there is missing data at prediction time
- ☐ Using imputed values results in higher accuracies than skipping data points or skipping features

7.

1 point

Consider data that has binary features (i.e. the feature values are 0 or 1) with some feature values of some data points missing. When learning the best feature split at a node, how would we best modify the decision tree learning algorithm to handle data points with missing values for a feature?



We choose to assign missing values to the branch of the tree (either the one with feature value equal to 0 or with feature value equal to 1) that minimizes classification error.



We assume missing data always has value 0.



We ignore all data points with missing values.