graphlab.nearest_neighbor_classifier.NearestNeighborClassifier.c

NearestNeighborClassifier. | classify | (dataset, max_neighbors=10, radius=None, verbose=True)

Return the predicted class for each observation in dataset, based on the closest neighbors stored in the nearest neighbors classifier model.

Parameters: dataset : SFrame

Dataset of new observations. Must include columns with the same names as the features used for model training, but does not require a target column. Additional columns are ignored.

verbose: bool, optional

If True, print progress updates.

max_neighbors: int, optional

Maximum number of neighbors to consider for each point.

radius: float, optional

Maximum distance from each point to a neighbor in the reference dataset.

Returns: out : SFrame

An SFrame with model predictions. The first column is the most likely class according to the model, and the second column is the predicted probability for that class.

① See also
create , predict , predict_topk

Notes

- If the 'radius' parameter is small, it is possible that a query point has no qualified neighbors in the training dataset. In this case, the resulting class and probability for that query are 'None' in the SFrame output by this method. If the target column in the training dataset has missing values, these predictions will be ambiguous.
- Ties between predicted classes are broken randomly.

Examples