

sklearn.datasets.make_circles

Used to generate points forming two circles. This function is ideal for visualizing and testing clustering and classification algorithms, as well as understanding non-linear data structures.

1. **n_samples** (int, default=100):
 - Sample Size
 - Description: Specifies the total number of points to generate, evenly split between the two circles.
 - Example: To simulate a scenario with 500 points, set `n_samples=500`.
2. **shuffle** (bool, default=True):
 - Randomization
 - Description: If set to True, the samples are shuffled, randomizing the order of the generated points between the circles.
 - Example: To keep the generated points organized (inner circle one color and outer circle another color), set `shuffle=False`.
3. **noise** (float or None, default=None):
 - Gaussian Noise
 - Description: Adds Gaussian noise to the samples, with the value representing the standard deviation of the noise. To do this, create a matrix with normally sampled values of `sd = noise` and then add it to the data matrix we have (which are (x,y) coordinates)
 - Example: To add small noise to the dataset, which could simulate measurement errors, set `noise=0.1`.
4. **random_state** (int, RandomState instance, or None, default=None):
 - Random Seed
 - Description: Determines the randomness of the dataset generation for reproducibility.
 - Example: To ensure the dataset can be regenerated in the same way, set `random_state=42`.
5. **factor** (float, default=0.8):
 - Inner Circle Scaling
 - Description: A scale factor between 0 and 1 that determines the size of the inner circle relative to the outer circle.
 - Example: To create an inner circle that is half the diameter of the outer circle, set `factor=0.5`.