

```
# impute missing values - axis=0: impute along columns
imputer = SimpleImputer(missing_values=np.nan, strategy='mean')
imputer.fit_transform(data)
'Output':
array([[5., 5., 8.],
       [9., 3., 5.],
       [8., 6., 4.],
       [5., 5., 2.],
       [2., 3., 9.],
       [5., 8., 7.],
       [1., 5., 5.]])
```

Generating Higher-Order Polynomial Features

Scikit-learn has a module called `PolynomialFeatures` for generating a new dataset containing high-order polynomial and interaction features based off the features in the original dataset. For example, if the original dataset has two dimensions $[a, b]$, the second-degree polynomial transformation of the features will result in $[1, a, b, a^2, ab, b^2]$.

```
# import packages
from sklearn.preprocessing import PolynomialFeatures

# create dataset
data = np.array([[5,8],[9,3],[8,6],
                 [5,2],[3,9],[8,7],
                 [1,5]])

data
'Output':
array([[5, 8],
       [9, 3],
       [8, 6],
       [5, 2],
       [3, 9],
       [8, 7],
       [1, 5]])
```

```
# create polynomial features
polynomial_features = PolynomialFeatures(2)
data = polynomial_features.fit_transform(data)
data
'Output':
array([[ 1.,  5.,  8., 25., 40., 64.],
       [ 1.,  9.,  3., 81., 27.,  9.],
       [ 1.,  8.,  6., 64., 48., 36.],
       [ 1.,  5.,  2., 25., 10.,  4.],
       [ 1.,  3.,  9.,  9., 27., 81.],
       [ 1.,  8.,  7., 64., 56., 49.],
       [ 1.,  1.,  5.,  1.,  5., 25.]])
```

Machine Learning Algorithms

This chapter provides an introduction to working with the Scikit-learn library for implementing machine learning algorithms.

In the next chapters, we'll implement supervised and unsupervised machine learning models using Scikit-learn. Scikit-learn provides a consistent set of methods, which are the **fit()** method for fitting models to the training dataset and the **predict()** method for using the fitted parameters to make a prediction on the test dataset. The examples are geared at explaining working with Scikit-learn; hence, we are not so keen on the performance of the model.