Preprocessing	package/link/comment	method/class(Parallel:○, Need label: *)
I/O	Pandas **The main method for reading/writing CSV, JSON, XML etc. files. It can parse the data type of columns automatically.**	• pandas.read_csv • pandas.read_sql_table • pandas.DataFrame.to_json • pandas.DataFrame.to_xml • etc.
	Numpy **The main method for saving a temporary n-dimensional array/tensor.**	• numpy.save • numpy.load
Normalization	scikit-learn	• sklearn.preprocessing.normalize
Encoding Categorical Features	scikit-learn **Convert categorical features to integer codes.**	• sklearn.preprocessing.OrdinalEncoder • sklearn.preprocessing.OneHotEncoder • etc.
	tensorflow **Convert categorical features to integer codes.**	○ tf.one_hot
Discretization	scikit-learn **Partition continuous features into discrete values.**	• sklearn.preprocessing.KBinsDiscretizer • sklearn.preprocessing.Binarizer • etc.
Working with Missing Data	Pandas **Detect/Fill/Drop the missing data.**	• pandas.DataFrame.fillna • pandas.DataFrame.dropna • pandas.DataFrame.interpolate • etc.
Noise Reduction	scikit-learn **Remove samples with noise by using clustering.**	• sklearn.cluster.KMeans • sklearn.cluster.AffinityPropagation • sklearn.cluster.DBSCAN • etc.
	Numpy/Scipy/Numpy **Remove the noise from a sequence by using the convolution operator(filter).**	• numpy.convolve • scipy.signal.convolve ○ tf.nn.conv1d
	scikit-learn **Remove the noise from a sequence by using the Regression.**	• sklearn.linear_model.LinearRegression
Data Reduction	scikit-learn **Reduce the number of attributes by using the SVD algorithm of matrix.**	• sklearn.decomposition.PCA • sklearn.decomposition.KernelPCA • etc.
	scikit-learn **Remove redundant attributes by correlation analysis.**	* scipy.stats.chisquare • etc.
Data Visualization		
Bar maps, curves, and pie charts	matplotlib **The most popular plotting library for the Python programming language.**	• matplotlib
model graph, histograms of training logs	tensorboard **TensorBoard base on TensorFlow provides the visualization and tooling needed for machine learning experimentation.**	• tensorflow/keras
Prediction		
Classification	scikit-learn **The main method for classification tasks in tabular data.**	* sklearn.tree.DecisionTreeClassifier
	tensorflow/keras **The main method for classification tasks in image datasets and audio datasets.**	⊛ tensorflow/keras • etc.
Regression	scikit-learn **Almost all popular regression algorithms are based on their classification version.**	* sklearn.linear_model.LinearRegression * sklearn.svm.SVR * sklearn.ensemble.RandomForestRegressor • etc.