**Pipeline:**

The execution of the workflow is in a pipe-like manner. Scikit-learn is a powerful tool for machine learning, provides a feature for handling such pipes under the sklearn.pipeline module called Pipeline. A machine learning pipeline can be created by putting together a sequence of steps involved in training a machine learning model.

It can be used to automate a machine learning workflow. The pipeline can involve pre-processing, feature selection, classification/regression, and post-processing.

**Process of Pipeline:**

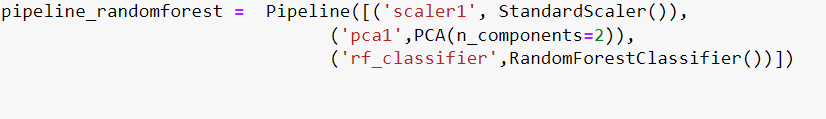
After loading the data, split it into training and testing then build pipeline object. wherein standardization is done using StandardScalar() and dimensionality reduction using PCA(principal component analysis) both of these with be fit and transformed(these are transformers), lastly the model to use is declared here it is Random forest and KNN, this is the estimator. The pipeline is fitted and the model performance score is determined.

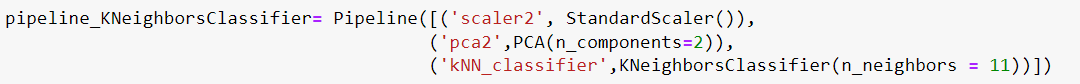
* **Pipelines creaton:**

1. First load the data and split it into training and testing.

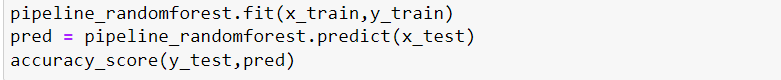


1. Data Preprocessing by using Standard Scaler that is scaled down the data which is based on normal distribution.
2. Reduce Dimension using PCA.
3. Apply Classifier, we used classification Algorithm such as Random Forest & KNN (with n\_component=11) and the code are following as shown in below.

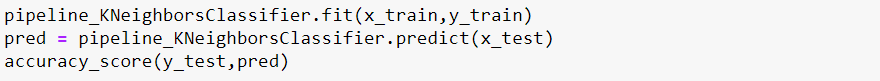




1. Let’s call fit() method of the pipe object on our training data and get the training and test scores for random forest and KNN at n=11 as shown in below.



Accuracy score for random forest is 0.96.



Accuracy score for KNN is 0.95

1. Conclusion is that we get good accuracy for random forest.

**Conclusion**

Pipelines keep our preprocessing steps and giving the summary of our models, making the machine learning workflow much easier. We can apply more than one preprocessing step if needed before fitting a model in the pipeline.

The main benefit for us has been being able to come back to a project and following the workflow I set with pipelines.