**Pipelines**

Estimator / algorithm – using fit() –> Transformer / Model

Example:

* From estimator to learned models
  + LogisticRegression -> LogisticRegressionModel
  + Pipeline -> PipelineModel
    - Assuming the pipeline includes at least one estimator
* Feature Transformers
* Tokenizer (is a transformer)
* HashingTF (is a transformer)

**Model selection and tuning**

Three required components for *CrossValidator()* and *TrainValidationSplit()*

1. Estimator: It can be a pipeline
2. ParamGridBuilder: *ParamGridBuilder()*
3. Evaluator: *RegressionEvaluator(), BinaryClassifictionEvaluator(), MulticlassClassficiationEvaluator()*

After fitting CrossValidator and TrainValidationSplit object, we will get a **learned model** with parameters that generate the best result.