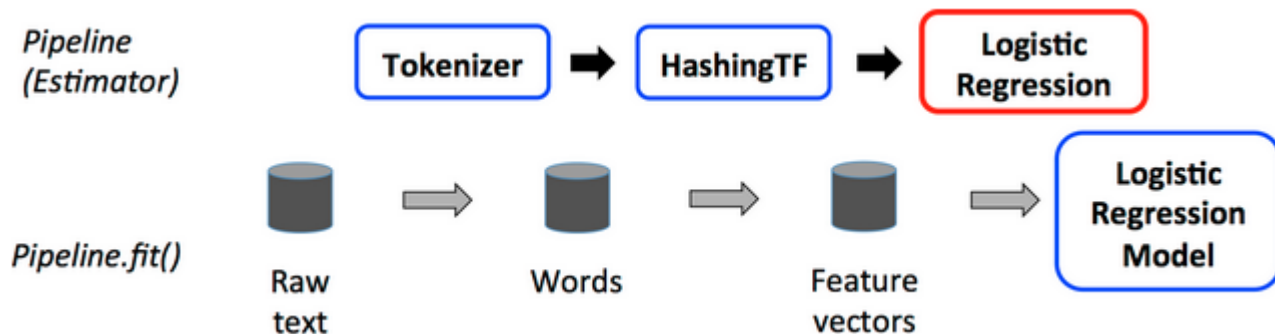# Pipelines

Pipelines

- Combine several algorithms into a single workflow
- Components:
    - DataFrames
    - Transformers: an algorithm (model) that transforms one DF into another (transform method)
    - Estimators: an algorithm that can be fit on a DF to produce a model, that is, a Transformer (fit and transform methods)
        - In the pipeline, all estimators become transformers
    - Parameters: common API for specifying parameters

Pipelines

- A Pipeline is an Estimator itself
  - when fit() method runs, every Estimator becomes a Transformer and the Pipeline outputs a model (a Transformer)

Pipelines

- A Pipeline is an Estimator itself
  - when the model's transform() method runs, each Transformer operates on the input DataFrame and produces a transformed DataFrame