

Raw data



Cleaner



Cleaned data



Vectorizer



Vectorized data



Dimensionality reduction



Reduced vectorized data



Estimator



Predictions