```
广义线性模型(Generalized Linear Models)
              线性二次判别分析(Linear and Quadratic Discriminant Analysis)
              核岭回归(Kernel ridge regression)
              支持向量机(Support Vector Machines)
              随机梯度下降(Stochastic Gradient Descent)
              最近邻(Nearest Neighbors)
              高斯过程(Gaussian Processes)
              交叉分解(Cross decomposition)
监督学习
              朴素贝叶斯(Naive Bayes)
              决策树(Decision Trees)
              集成方法(Ensemble methods)
              多类和多标签算法(Multiclass and multilabel algorithms)
              特征选择(Feature selection)
              半监督(Semi-Supervised)
              保序回归(Isotonic regression)
              概率校准(Probability calibration)
              神经网络模型(监督)(Neural network models (supervised))
                高斯混合模型(Gaussian mixture models)
                流形学习(Manifold learning)
                聚类 (Clustering)
                双聚类(Biclustering)
无监督学习
                分解信号分量(矩阵分解问题)(Decomposing signals in components (matrix factorization problems))
                协方差估计(Covariance estimation)
                异常值检测(Novelty and Outlier Detection)
                密度估计(Density Estimation)
                神经网络模型(无监督)(Neural network models (unsupervised))
                   交叉验证:评估估计器性能(Cross-validation: evaluating estimator performance)
                   调整估计器的超参数(Tuning the hyper-parameters of an estimator)
                   模型评价:量化预测的质量(Model evaluation: quantifying the quality of predictions)
模型选择和评估
                   模型持久性( Model persistence)
                   验证曲线: 绘制分数以评估模型 ( Validation curves: plotting scores to evaluate models )
                 管道和特征:组合估计量
                 特征提取
                 预处理数据
                 无监督隆维
数据集转换
                 随机投影
                 内核近似
                 成对度量,亲和度和内核
                 变换预测目标
                       通用数据集API(General dataset API)
                      玩具数据集(Toy datasets)
                       示例图像(Sample images)
                       - 样品发生器(Sample generators)
                       symlight / libsym格式的数据集(Datasets in symlight / libsym format)
                       从外部数据集加载(Loading from external datasets)
                       Olivetti面数据集(The Olivetti faces dataset)
                       20个新闻组文本数据集(The 20 newsgroups text dataset)
                       从mldata.org存储库下载数据集(Downloading datasets from the mldata.org repository)
数据集加载实用程序
                       野生面部识别数据集中的标记面(The Labeled Faces in the Wild face recognition dataset)
                       森林covertypes (Forest covertypes)
                       RCV1数据集(RCV1 dataset)
                       波士顿房价数据集(Boston House Prices dataset)
                       乳腺癌威斯康星(诊断)数据库(Breast Cancer Wisconsin (Diagnostic) Database)
                       糖尿病数据集(Diabetes dataset)
                       手写数字数据集的光学识别(Optical Recognition of Handwritten Digits Data Set)
                       虹膜植物数据库(Iris Plants Database)
                       Linnerrud数据集(Linnerrud dataset)
计算规模的策略: 更大的数据 -
                              一使用核外学习来扩展实例(Scaling with instances using out-of-core learning)
              预测延迟(Prediction Latency)
计算性能
              预测吞吐量(Prediction Throughput)
              技巧和窍门(Tips and Tricks)
```

sklearn