1. 监督学习及相关模型

监督学习

常用模型 / 应用场景	模型名称		API
分类	LogisticRegression(逻辑回归)		from sklearn.linear_model import LogisticRegression
	SGDClassifier(随机梯度下降分类器)		from sklearn.linear_model import SGDClassifier
	SVC(支持向量机分类器)		from sklearn.svm import LinearSVC
	Naive Bayes(朴素贝叶斯)		from sklearn.naive_bayes import MultinomialNB
	KNN(K 近邻分类器)		from sklearn.neighbors import **KNeighborsClassifier**
	DecisionTree(决策树)		from sklearn.tree import DecisionTreeClassifier
	Ensemble (集成分类模型)	RandomForestClassifier (随机森林)	from sklearn.ensemble import RandomForestClassifier
		GradientBoostingClassifier (梯度提升决策树)	from sklearn.ensemble import GradientBoostingClassifier
回归	LinearRegression(线性回归)		from sklearn.linear_model import LinearRegression
	SGDRegressor(随机梯度下降回归器)		from sklearn.linear_model import SGDRegressor
	SVR(支持向量机回归器)		from sklearn.svm import SVR
	KNN(K 近邻回归器)		from sklearn.neighbors import **KNeighborRegressor**
	DecisionTreeRegressor(回归树)		from sklearn.tree import DecisionTreeRegressor
	Ensemble (集成回归模型)	RandomForestRegressor (随机回归森林)	from sklearn.ensemble import RandomForestRegressor
		ExtraTreesRegressor (极端回归森林)	from sklearn.ensemble import ExtraTreesRegressor
		GradientBoostingRegressor (梯度提升回归树)	from sklearn.ensemble import GradientBoostingRegressor

性能评估

场景	评估方法	API
分类	Accuracy(准确度)	estimator.score(X_test,y_test)
	精确率、召回率、F1 分数	from sklearn.metrics import Classification_report
回归	R_squared(模型自带)	estimator.score(X_test,y_test)
	R_squared, MSE(平方相对误差), MAE(平方绝对误差)	from sklearn.metrics import r2_score, mean_squared_error, mean_absolute_error

2. 无监督学习及相关模型

无监督学习

场景	常用模型	API
聚类	K-Means(K 均值算法)	from sklearn.cluster import KMeans
降维	PCA(主成分分析)	From sklearn.decomposition import PCA

性能评估

场景	评估方法	API
聚类	ARI(Adjust Rand Index)	from sklearn.metrics import adjust_rand_score
	Silhouette Coefficient(轮廓系数) Separation(分离度)	from sklearn.metrics import silhouette_score
降维	_____	

3. 特征提取及模型选择

特征提取

场景	常用模型		API
特征提升	DictVectorizer(文本特征向量化)		from sklearn.feature_extraction import DictVectorizer
	词袋法	CountVectorizer	from sklearn.feature_extraction.text import CountVectorizer
		TfidfVectorizer	from sklearn.feature_extraction.text import TfidfVectorizer
特征筛选	feature_selection		from sklearn import feature_selection

模型选择

场景	常用模型	API
模型正则化	Lasso(L1 范数正则化)	from sklearn.linear_model import Lasso
	Ridge(L2 范数正则化)	from sklearn.linear_model import Ridge
模型校验	留一法 、交叉验证法	
超参数搜索	GridSearch(网格搜索)	from sklearn.grid_search import GridSearchCV
简化系统搭建	Pipeline	from sklearn.pipeline import Pipeline