

机器学习

机器学习概述

人工智能概述

什么是机器学习

数据

数据集构成

特征值和目标值

模型

预测

机器学习开发流程

获取数据

数据处理

特征工程

机器学习算法训练-模型

模型评估

应用

机器学习算法分类

监督学习（有目标值）

分类算法

回归算法

无监督学习（无目标值）

聚类

机器学习

特征工程

数据集

可用数据集

sklearn数据集

load\_\*小规模数据集

fetch\_\*大规模数据集

Bunch类型

数据集划分

model\_selection.train\_split()

特征抽取feature\_extraction

字典特征抽取

DicVectorizer

sparse矩阵-节省空间

文本特征抽取

CountVectorizer

TfidfVectorizer

特征预处理

无量纲化

归一化MinMaxScale()

标准化StandardScale()

特征降维

特征选择

过滤式

删除低方差特征VarianceThreshold()

相关系数

取值范围(-1,1)

嵌入式

主成分分析（PCA降维）