



中国石油大学
CHINA UNIVERSITY OF PETROLEUM

《油气人工智能基础及应用》

3.1 常规分类方法

董少群

dshaoqun@163.com

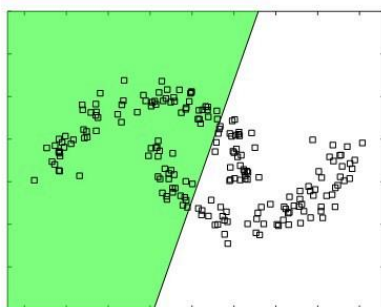
理学院数学系

CUP

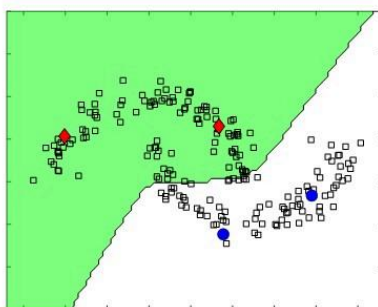
厚积薄发 开物成务



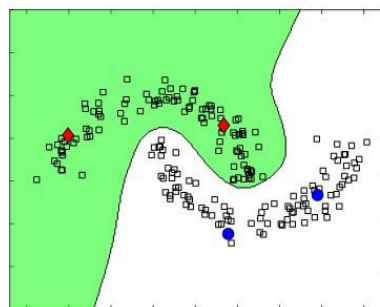
中国石油大学
CHINA UNIVERSITY OF PETROLEUM



无监督聚类



有监督分类

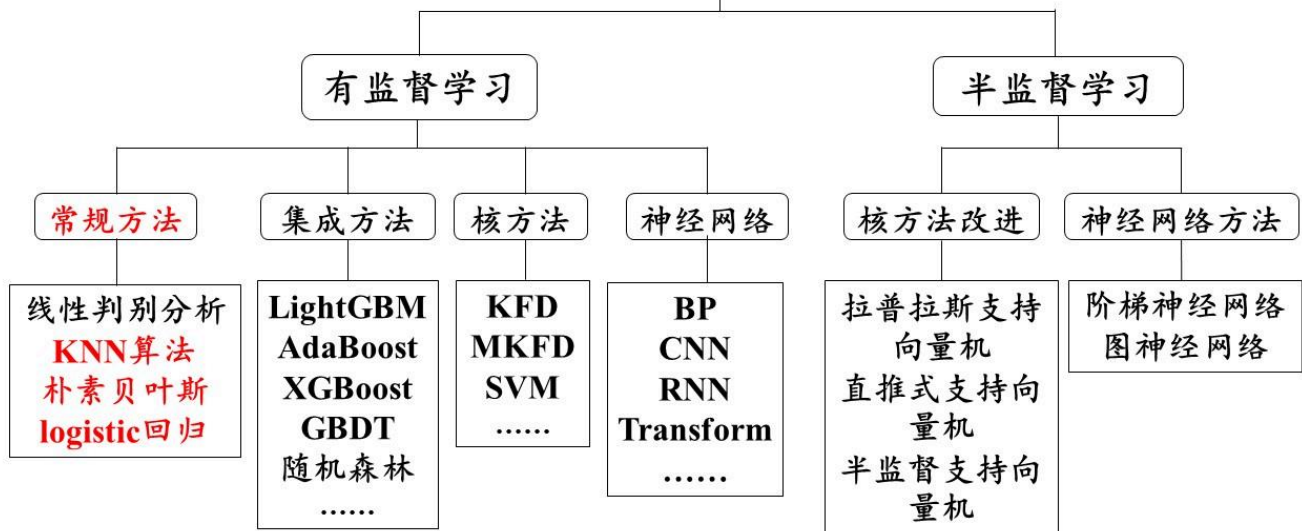


半监督分类

[1] Shaoqun Dong, et al. Fracture identification by semi-supervised learning using conventional logs in tight sandstones of Ordos Basin, China[J]. Journal of Natural Gas Science and Engineering. 2020, 76: 103131.

2

机器学习中的分类方法



常规分类方法介绍

- 一、KNN (K nearest neighbors)
- 二、朴素贝叶斯 (Navie Bayes)
- 三、逻辑回归 (Logistic regression)

一、KNN算法

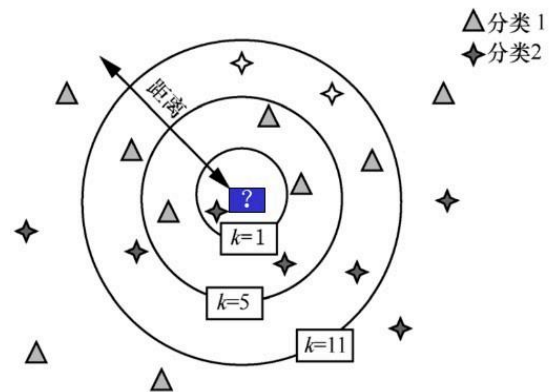


中国石油大学
CHINA UNIVERSITY OF PETROLEUM

1. KNN概述

1968年，KNN算法由Cover& Hart提出。是通过测量不同特征值之间的**距离**进行分类的算法。

它的思路是：如果一个样本的 k 个最近邻的样本的**大多数**属于某一个类别，则该样本也属于这个类。



5

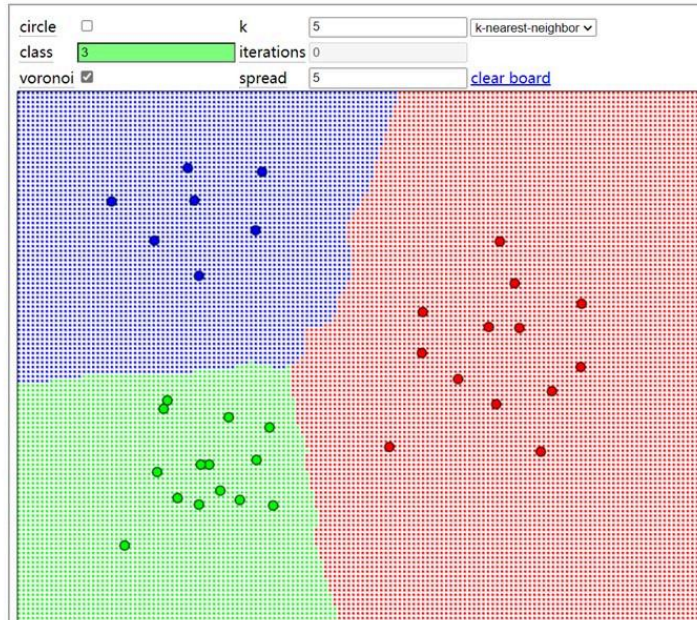
一、KNN算法



中国石油大学
CHINA UNIVERSITY OF PETROLEUM

<https://martin-thoma.com/k-nearest-neighbor-classification-interactive-example/>

<https://lecture-demo.ira.uka.de/knn-demo/>



6

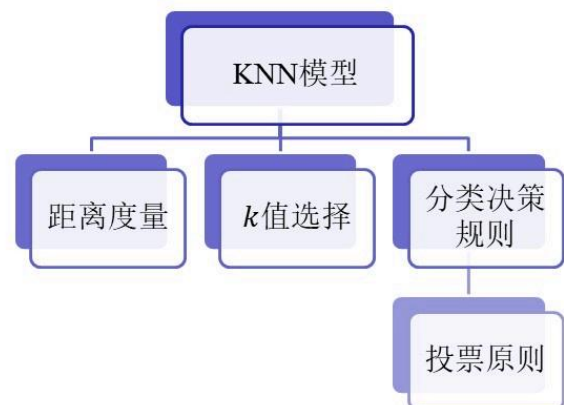
一、KNN算法



中国石油大学
CHINA UNIVERSITY OF PETROLEUM

2. KNN 原理—三个关键参数

- ①对于输入样本，依据给定的距离度量方式（一般使用欧式距离）
- ②选择合适的 k 值（交叉验证）
- ③在样本集中找到实例最近邻的 k 个样例，通过 k 个最近邻样例的类别表决出新实例的类别（多数表决）。



7