

浮动主题

冗余计算

NMI

同质化

SU

分子IG

分母H(X)

数值上等于MI

公式: $H(X) - H(X|Y)$ $H(X)$ 的计算

直接代入数据

数据离散化 (后4位/MDLP) 后代入

X: 离散; Y: 连续特征

X, Y: 连续特征

连续数据

离散数据 (类标签)

离散数据

kde

频率当概率

MDLP离散方法

1. 将连续数据直接看成离散数据

忽略数据性质

2. 连续数据处理

计算不严谨, 分子分母不一致的处理方式

3.2 离散化处理

离散化后再分别计算分子分母

3.1 均连续化 (kde) 计算

