FuzzyID3思路：

总体结构图：

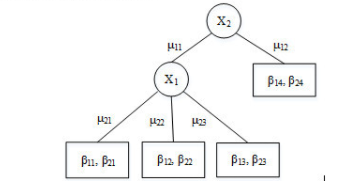


图1

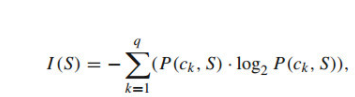
1. 数据集【以两组数据为例】{问题1：如何获得足够的可靠的数据集？？}
   1. 格式：脉搏、眨眼频率、用户状态
   2. 示例：
      1. 【70、0.35、安全】
      2. 【80、0.3、疲劳】
      3. 【65、0.1、疲劳】
2. 隶属函数🡪【用来将原始连续数据模糊化（概率化）】{问题2：如何确定隶属函数？？}
   1. 示例：Y1 = F（脉搏数、快慢）、Y2 = F（眨眼频率，快慢）
   2. 示例：Y3 = F（脉搏数，P(安全)）、Y4 = F（眨眼频率，P(安全)）
3. 数据处理：🡪将原始连续数据模糊化（概率化）【使用隶属函数】

【P(脉搏快)，P(脉博慢)，P(眨眼快)，P(眨眼慢)，安全/疲劳】

以处理第一个数据为例：【0.3，0.7，0.4，0.6，安全/疲劳】

1. 计算特征属性值的模糊熵【根据大小决定第n次分类使用何种特征】--- G（脉搏快）
   1. 

S：总体样本集合、Xj：特征X的属性值为j【脉搏、快】



P（ck , S）：S中某一类【安全/疲劳】样本的隶属度的和 / 集合总隶属度的和

【问题3：所有特征都算 / 只对待算的特征（如本例中 脉搏】求和？】





Sjv：S 中 第v个 特征X的属性值为j 的 样本

SjFv：S 中 特征X的属性值j的隶属度为Fv的样本的集合

| SjFv |： S 中 特征X的属性值j的隶属度为Fv的样本的隶属度的和

I(SjFv)：

j—快慢 ， X—脉搏 / 眨眼

例：

数据集 [0.2, 0.8, 0.3, 0.7，安全 ] [0.1, 0.9, 0.15, 0.85，疲劳]

计算：

G(脉搏快 , S) = I(S) – E(脉搏快 , S)

设X = (0.2) / (0.2+0.1) 、Y = (0.1) / (0.2+0.1)

I(S) = - ( X\*log₂X) – (Y\*log₂Y)

1. 计算特征属性值的权重：【决定特征的属性值下有无子叶节点】【图1 μ11等】

W（脉搏快） = G（脉搏快） /（ G（脉搏快） + G（脉搏快））

1. 根据根节点修改数据集

例：G（脉搏慢）> G（脉搏快）

将原始数据中所有脉搏慢项取出并删去脉搏特征，以此构成新的数据子集。

1. 用新子集递归 5.6 点 直至所有特征被删掉，构建出决策树。

例：脉搏的模糊熵大，脉搏快的权重大于脉博慢。

总体样本

（脉搏慢） 安全，疲劳（脉搏快）

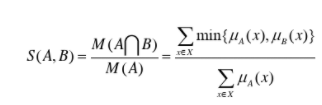
安全，疲劳（眨眼快）安全，疲劳（眨眼慢）

1. 根据 决策树 计算 模糊规则【即树中某条路径】 的真实度S： 【图1 β11等】

S（脉搏快，安全）、S（脉搏快，疲劳）

S’（眨眼快，安全）、S’（眨眼慢，安全）、S’（眨眼快，疲劳）、S’（眨眼慢、疲劳）

S’：由脉搏慢构成的新的子集



A：脉搏快、B安全

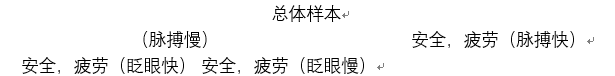
X：数据集的所有样本 x：数据集中某个样本

μA(x)：样本脉搏数对应 脉搏快 的概率。【隶属函数 Y1】

μB(x)：样本脉搏数对应 用户安全 的概率。【隶属函数Y3】

---------------------------------------------模糊决策树构建完成---------------------------------------

1. 计算总体隶属度：【输入样本时】



首先计算最右边路径的Upath

根据输入样本计算出

脉搏快的概率P1

对于该路径模糊规则的真实度S1【脉搏快，安全】 S2【脉搏快，疲劳】

Upath1 = P1 \* S1\* μ1 Upath2 = P1 \* S2 \* μ2

二者取大。

其次计算最左边路径的 Upath

根据输入样本计算出

脉搏慢的概率P2、眨眼快的概率P3

对于该路径模糊规则的真实度S3【眨眼快，安全】 S4【眨眼慢，疲劳】

{问题4，此处不考虑上层的脉搏慢？}

Upath3 = P2 \* S3 Upath4 = P3 \* S4

二者取大。

同理计算第三条路径，最终Upath 大者为分类结果。