

# 第一章 绪论

Author: 中山大学 17数据科学与计算机学院 YSY

<https://github.com/ysyisyourbrother>

## 引入

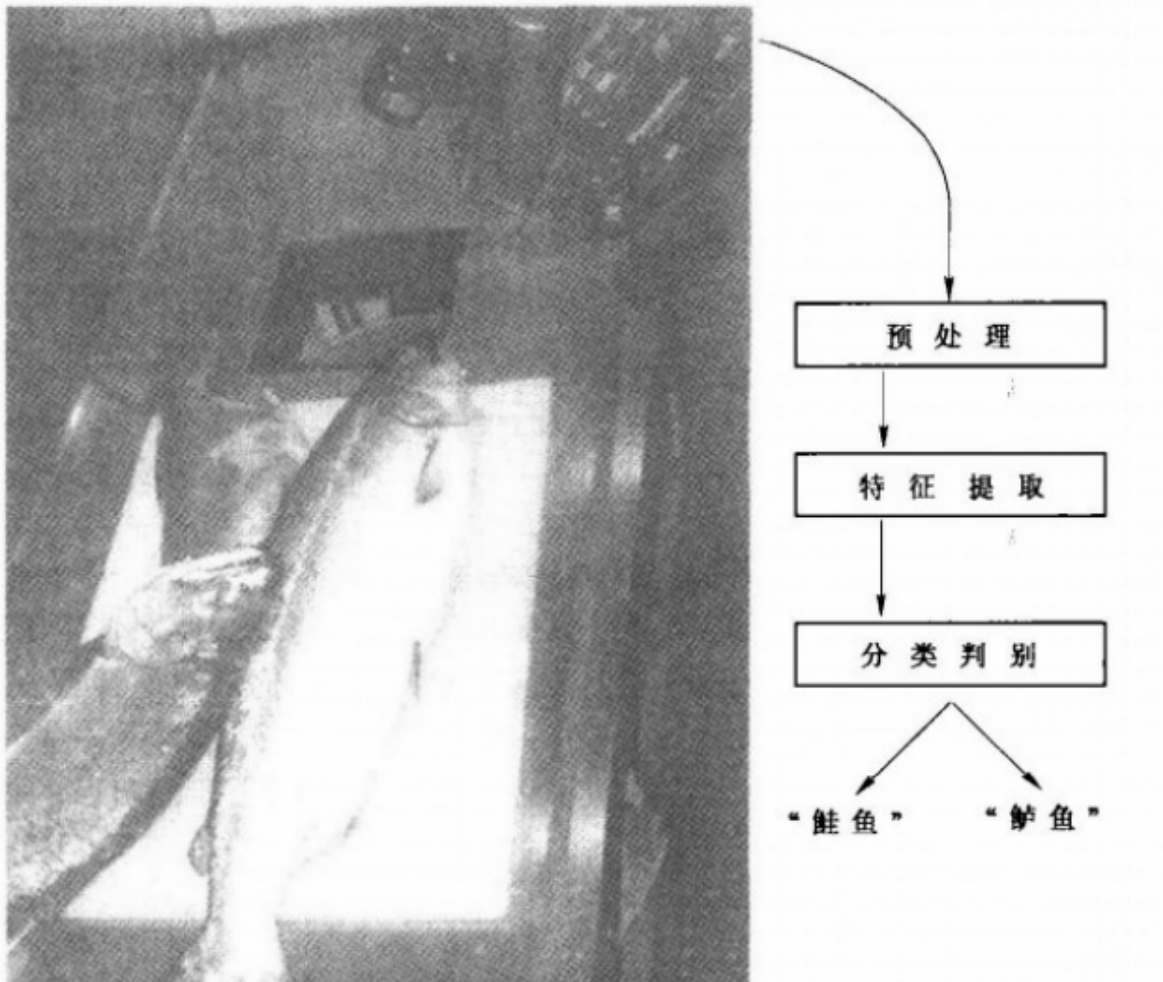
如果要区分鲑鱼和鲈鱼，我们先提取他们的**特征**

如果鲑鱼和鲈鱼之间确实存在某种差异，我们称之为不同的模型，即可以用数学形式表达的不同特征的描述。

模式分类的最终目的和处理方法就是：

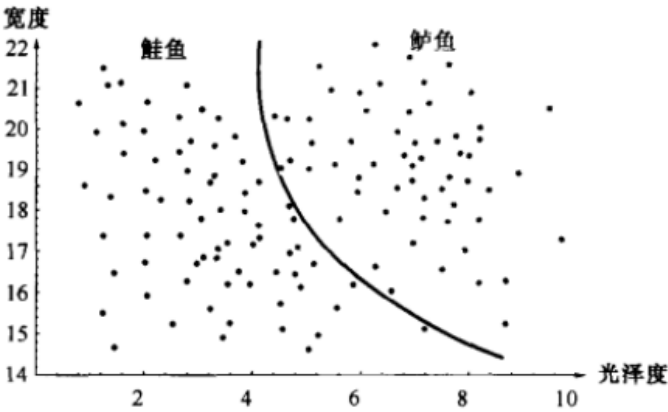
1. 将模型分成几类，对感知到的数据进行处理，去除干扰
2. 选择出与感知数据最接近的模型类别

本质就是根据模型的特征对数据进行分类。



我们真正的任务是要确定一个决策，使代价函数最小。

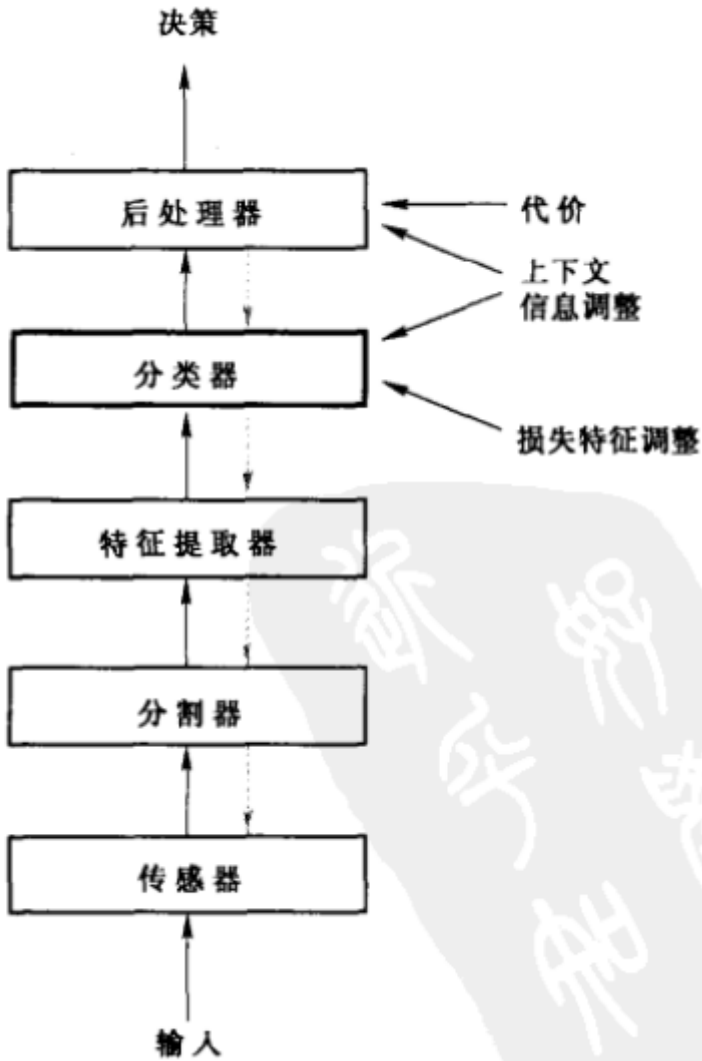
图 1-6 图中标示出的判决曲线是对训练样本的分类性能和分界面复杂度的一个最优折中。因而对将来的新模式的分类性能也很好



## 模式识别系统

模式分类的三种基本操作是：预处理，特征提取，分类判别。

图 1-7 很多模式识别系统都可以按这种方式划分为模块。传感器把图像声音等物理输入转换为输入信号。分割器将物体与背景及其他物体分开。特征提取器测量用于分类的物体属性。分类器根据特征给物体赋予类别标记。最后，后处理器作一些其他考虑，比如上下文信息、错误代价、选择合适的动作。尽管这种描述强调了信息单方向自下而上流动，但是有些系统采用了反馈机制(图中向下的浅色线)



## 学习和适应

### 有监督学习

存在一个教师信号，对训练样本集中的每个输入提供类别和分类代价，并寻找降低总体代价的方向。

### 无监督学习

没有显示的教师。系统对输入样本自动聚类或自然组织

## **强化学习**

不需要指明类别和教师信号，只需要对这次分类任务给出对或者错的反馈，然后机器就会去学习。