**《人工智能与模式识别》平时作业**

**第一章作业**：(共5分)（简答题 ×3）

1. 简述模式识别的目标和原理。(1分)
2. 简述设计一个完整的模式识别系统需要有哪些环节？(2分)
3. 超市的商品上贴有条形码，结账时收银员通过扫描条形码获取商品信息，在此过程中的条形码识别属于模式识别问题吗？请说明原因。(2分)

**答案：**

1. 答：模式识别的目标：寻找一种区分不同类的模式的“方法”，称为“分类方法”，该方法可以通过计算机技术来实现。给定“新”的模式，利用找到的分类方法将其分配给最接近的模式类。

模式识别原理：

（1）思想：从过去的经验中学习。

（2）训练集：搜集的历史数据，用于观察数据中的规律。

（3）特征提取：找到一种新的数据表示方法，使得规律更容易发现。

（4）模型学习：学习规律的过程。

（5）模型：从训练集中学习到的规律的体现，其表现形式多种多样。

（6）测试集：为了评估学到的规律而搜集的数据。

（7）预测/测试：将规律应用到新实例上，得到预测结果。

2. **答：**包括下面环节：

模式采集：借助于物理设备，如传感器、摄像头、变送器等，采集和存储模式。

预处理：通过各种滤波降噪措施，降低模式采集过程的干扰因素，增强有用的信息。

特征提取：提取对识别有用的特征。提取的特征类型因问题而异。

分类器的训练/学习：首先根据训练数据特点 (规模、稀疏性、特征冗余性、决策边界复杂性等)，选择分类器类型，然后利用训练数据集自主学习分类器所需的参数，获得训练好的分类器。

3. **答：**条形码或称条码（barcode）是将反射率相差很大的黑条（简称条）和白条（简称空）按照一定的**编码**规则组合起来，用以代表一定的字母、数字等信息。在进行辨识时，使用条码阅读机扫描，得到一组反射光信号，此信号经光电转换后变为一组与黑条和白条相对应的电子讯号，经**解码**后还原为相应的字母或数字，再传入电脑。该辨识过程不是基于learning from experience，所以**不属于**模式识别问题。

**第二章作业**：(共5分)

（单选（0.5分）×2，简答题 ×2）

1.【单选】 关于分类模型的复杂度与泛化能力的关系，下列描述正确的是（ C ）

A. 分类模型越简单，模型的泛化能力越好

B. 分类模型越复杂，模型的泛化能力越好

C. 为提高分类模型的泛化能力，模型既不能太简单也不能太复杂

D. 简单的分类模型会产生更好的泛化能力

2. 【单选】关于模型的泛化性能的论述中，( B ) 是正确的论述。

A．模型在训练数据集上的训练误差越小，表明其泛化性能越好

B. 使用更多的训练数据可以提高模型的泛化性能

C. 训练集大小不变时，使用复杂的模型一定能提高模型的泛化性能

D. 模型的泛化性能定义为模型在测试数据集上的性能

1. 简述什么是“过拟合”现象，它与模型的泛化能力有什么关系？(1分)

4. 详述设计一个车牌识别系统的整个流程。(3分)

1. 答：由于过分追求对训练集中实例的正确分类，导致分类器的泛化能力降低，称为分类器训练过程中的“过拟合”现象。发生过拟合的模型，其泛化能力一定差。
2. 整个流程如下：

(1) 图像预处理：对采集的图像进行灰度化、图像去噪等处理，目的是突出车牌的主要特征，更好地提取车牌区域。

(2) 车牌区域定位：利用经典的车牌定位方法，从整幅图像中精准地找出包含车牌的区域，将该区域从原图像中分割出来。

(3) 车牌矫正：摄像头与车牌之间往往有着一定的角度，使得车牌图片呈非水平形态，为了精准识别，必须对非水平态的车牌图像进行矫正。

(4) 字符分割：旨在将定位好的车牌图像区域分割成单个字符。

(5) 字符识别和车牌显示：采用模板匹配法或基于机器学习的方法，对分割得到的单个字符进行识别，并将字符连接成字符串进行展示。