

---

# 人工智能导论课程报告

学号：2016202152

姓名：张雅坤

学院：信息学院

年级：2016 级本科

## 一、概述

本学期的课程学习主要分为两个部分，理论和实践。理论部分，主要讲授了搜索问题与算法，机器学习，深度学习，基于 TensorFlow 的深度学习实践项目；实践部分，则是动手实现智能小车。我在学习课程的过程中，对人工智能领域有了大致的了解，学习了相关算法和模型，同时在课程项目中接触到了许多硬件知识，在学习理论知识之余提高了自己的动手能力。下面我将从课堂学习和课程项目两个方面详细总结本学期的收获。

## 二、课堂学习

在理论学习部分，课程大致按照时间顺序，依次介绍了传统的实现人工智能的搜索算法、传统的机器学习算法、以及近几年比较热门的深度学习和神经网络方法。搜索算法部分主要学习了，在深度优先搜索和广度优先搜索的基础上从不同角度改进的分支限界法、A\*算法、alpha-beta 剪枝方法等等；机器学习部分主要学习了逻辑回归和线性回归模型、决策树模型、支持向量机模型、聚类算法、以及集成学习方法随机森林模型等；深度学习部分，主要学习了人工神经网络、卷积神经网络、循环神经网络、自动编码器等。其中，我印象比较深刻的模型有逻辑回归模型、决策树模型、高斯混合聚类模型。

逻辑回归模型主要应用于二分类任务，在二分类任务中，我们希望找到一个单调可微函数，使得

$$y = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases}$$

其中 $z = w^T x + b$ 为预测值， $y = 1$ 表示正例， $y = 0$ 表示反例。而对数几率函数恰恰符合该性质，

$$y = \frac{1}{1 + e^{-z}}$$

对该函数稍加变形，可以得出

$$P\{y = 1|x\} = \frac{e^{w^T x + b}}{1 + e^{w^T x + b}}$$
$$P\{y = 0|x\} = \frac{1}{1 + e^{w^T x + b}}$$

在进行分类时，对于给定的 $x$ 按上述二式计算并比较，最后将实例 $x$ 分类到二者之中较大的那一类。其中，该模型的参数为 $w^T$ ，在训练模型时首先进行初始化，然后使用最大似

---

然估计法进行多次迭代，直到收敛到阈值为止。

决策树模型包含一个根结点，若干个内部节点和若干叶节点，每个叶节点对应决策结果，每个非叶节点对应属性测试。决策树学习的目的是产生一棵泛化能力强，即处理未见示例能力强的决策树，建树的过程主要采取分治策略，其中最重要的步骤是划分最优属性集，为了使得决策树的分支节点所包含的样本尽可能属于同一类，可以选用信息增益、增益率、基尼指数作为选择划分属性的依据。由于决策树很容易出现过拟合现象，即模型“记住”所有的数据包括噪声但并没有从中“学到”分类规则，因此采用预剪枝和后剪枝的方法提升模型的泛化能力。

高斯混合聚类模型假设数据点呈高斯分布，训练时首先给出簇的数量并初始化每个簇的高斯分布参数，即均值和方差，然后计算每个数据点属于各个簇的后验概率（E-step），根据这些概率计算并更新高斯分布的参数使得数据点的概率最大化（M-step），重复迭代若干次直到参数收敛到阈值为止。

### 三、 课程项目

在课程项目部分，我们总共完成了三次展示，下面将依次详细说明我在项目中学到的知识和遇到的问题。

#### 1. 第一次展示

在第一次展示，我们完成了 Arduino 小车的硬件组装；实现了小车的基本动作，包括前进、停止、左转、右转、后退；实现了三大基础功能，包括循迹（红外循迹）、避障（红外避障与超声波避障）、遥控（手机蓝牙遥控）；此外，我们还实现了追踪特定物体功能，通过“迹”app 识别屏幕中特定颜色的物体，并通过蓝牙串口返回物体坐标值，程序根据坐标判断前进后退的方向和转角的大小。我主要负责小车的硬件组装，基本动作和基本功能的实现。

我主要依据网络上的视频讲解一步一步安装好小车，在这个过程中，我发现好多知识与我在计算机组成原理课程和数字逻辑与数字电路课程上学到的知识相呼应，仔细观察电路板，就可以发向上面标注了各种引脚编号和引脚类型，对应连接不同的元器件以实现多种功能。

安装小车完毕，我们接下来使用 Arduino 的 IDE 作为开发环境，通过其自带的 pinMode 函数初始化引脚，使用 digitalWrite 函数控制向某引脚写入高低电平，进而实现了控制电路板上灯的亮灭、控制小车轮子转向等，使用 digitalRead 函数获取传感器

---

的状态信息，使用 `analogWrite` 函数控制小车轮子的转速，这些基础函数是我们项目的基础。红外循迹功能的原理是，小车下面的左右循迹传感器在白色区域有信号，调用 `digitalRead` 函数返回 LOW，在黑色区域无信号，调用 `digitalRead` 函数返回 HIGH，根据两个传感器的返回值，可以判断小车是否偏离黑线（黑线宽度与两个传感器间距相等），如果偏离则可以进行对应的调节，进而实现循迹（循黑线）功能。红外避障的原理是，小车上方的左右避障传感器在有障碍是有信号，调用 `digitalRead` 函数返回 LOW，在无障碍时无信号，调用 `digitalRead` 函数返回 HIGH，根据两个传感器的返回值，可以判断小车是否接近障碍物，并可以采取对应的措施进行躲避，进而实现红外避障功能。超声波避障的原理是，通过对电压电平的控制可以控制超声波舵机云台的转动，进而使得小车可以探测前方、左侧、右侧的障碍物距离，并可以采取对应的措施进行躲避，进而实现超声波避障功能。蓝牙遥控的原理是，首先编写程序，使用蓝牙相关头文件 `SoftwareSerial.h` 中的 `Serial.read` 函数接收手机蓝牙向小车蓝牙发送的字符，根据字符判断指令并执行，在手机上安装“SPP 蓝牙串口”APP，在蓝牙连接完毕后，即可通过 APP 界面控制向小车发送字符，控制小车的行动。

在这一阶段中，我们主要遇到了三个问题。第一，最开始我们购买了简易版小车，需要自己动手焊接电路，由于在宿舍、教室不方便使用电烙铁，以及没有清楚了解电路板的引脚具体接线方法，我们最后选择另购插线版的小车；第二，对超声波舵机云台转向控制不熟练，不能很好的通过电压电平控制转向的角度和时间；第三，在实现追踪特定物体功能的过程中，一开始没能控制好转向的角度，使得追踪目标很容易因为小车转动角度过大而被甩出镜头，丢失追踪目标，然后进入盲目旋转模式失去控制，最后我们修改了小车每次偏转的角度。

## 2. 第二次展示

在第二次展示，我们实现了图像识别和语音控制两部分功能。我主要负责小车语音控制功能。

小车的语音控制主要流程是，使用 python 中 `pyaudio` 库进行 1.5s 的录音并存储为 .wav 文件，对该音频进行语音识别（匹配），识别出是“前”、“后”、“左”、“右”、“停”五个指令中的哪一个后，使用 `serial` 库中的函数将对应字符通过电脑蓝牙发送给小车蓝牙，从而控制小车的运动。

我实现了两种语音识别的方法，一种基于 DTW，一种基于 GMM-HMM 识别。基于 DTW

---

（动态时间规划）的原理是，首先录制“前”、“后”、“左”、“右”、“停”五个指令的语音数据集，对每个音频加窗分帧提取 MFCC（梅尔倒谱系数）特征，然后对同一指令的多个语音做 DTW 匹配，对应上的特征进行平均得到语音模板并存储下来，每次进行识别时，将待识别语音提取 MFCC 特征，将其与五个指令的模板分别做 DTW 匹配，将待识别语音识别为与之距离最近的指令。基于 GMM-HMM 识别的原理是，首先录制“前”、“后”、“左”、“右”、“停”五个指令的语音数据集，对每个音频加窗分帧提取 MFCC（梅尔倒谱系数）特征，将其和对应的标签输入进 GMM-HMM 模型中进行模型训练，每次进行识别时，将待识别语音提取 MFCC 特征，将其输入进 GMM-HMM 模型中，将待识别语音识别为概率最大的指令。

在这一阶段中，我们主要遇到的问题是语音识别效果受环境影响显著，在加有噪音好环境，识别的准确率小于无噪音环境。

### 3. 第三次展示

在第三次展示，我们实现了图像识别道路行进和交通标识两部分功能。我主要参与了小车道行进功能。

小车的道路行进功能主要流程是，首先搭好道路，将手机固定在小车上，使用 ip 摄像头将手机拍摄的图片传到电脑上，用电脑按键控制向小车发送相应蓝牙字符控制小车行动，记录下小车在道路上的视角图片（手机摄像头拍摄）和对应的电脑按键信息，构成数据集对 ANN（人工神经网络）进行训练，实验时，通过 ip 摄像头获取小车视角图片，对图片进行识别，判断应进行的操作，将对应的字符通过电脑蓝牙发送到小车蓝牙上控制小车行动。

在这一阶段中，我们主要遇到了三个问题。第一，数据集采集不够，模型识别不够准确；第二，数据采集不平衡，有很多的直行视角图片，转弯视角图片较少，使得模型对转弯的识别效果非常差；第三，延迟严重，通过 ip 摄像头将手机摄像头图片发送给电脑耗费很长时间，造成延迟，十分影响实验效果，我们考虑过将模型移植到手机上运行，但由于编写 Android APP 需要申请摄像头和蓝牙权限等操作过于复杂，最终没能实现。

## 四、 总结

通过人工智能导论这门课，我了解到了很多关于人工智能领域的知识，对于知识体系有了一个整体的把握，同时，我在动手实践的过程中学习到很多硬件知识，比如蓝牙串口知识，

---

提升了自己的动手实践能力,也体会到将理论应用到实际解决生活中各种真实复杂的问题是一件不容易的事,因为现实中没有绝对严格的条件假设,也很难有绝对完美的结果,而是蕴含着各种灵活多变的情景、复杂繁多的可能性、不理想不精确的条件,以及无法求得最优解而不得不退而求其次的妥协。

