

# 人工智能简介作业答案

2025 年 5 月 20 日

## 1.1 人工智能简介

**理论：找出 5 个不同的应用场景，分析其输入、输出、反馈分别是什么？**

### 1. 智能客服机器人

- 输入：用户的文本或语音查询（例如：“我的订单到哪里了？”）。
- 输出：机器人提供的文本或语音回答、相关信息链接或执行的操作（例如：“您的订单已于 X 月 X 日发货，预计 X 月 X 日送达。”）。
- 反馈：用户对回答的满意度评价（例如：点赞/点踩、追问、转人工服务）。系统根据反馈调整回答策略或知识库。

### 2. 自动驾驶汽车

- 输入：来自各种传感器的数据（例如：摄像头捕捉的图像、雷达和激光雷达探测到的障碍物距离和速度、GPS 定位信息、高精度地图数据）。
- 输出：对汽车的控制指令（例如：加速、减速、转向、刹车）。
- 反馈：汽车实际的行驶状态（例如：速度、方向、与障碍物的距离）、乘客的舒适度、是否发生碰撞或违章。系统根据反馈调整驾驶策略和模型参数。

### 3. 个性化推荐系统（例如：电商、视频网站）

- 输入：用户的历史行为数据（例如：浏览记录、购买记录、搜索记录、点赞/收藏）、用户画像（例如：年龄、性别、兴趣标签）、物品信息（例如：商品属性、视频标签）。
- 输出：向用户展示的推荐列表（例如：推荐商品、推荐视频）。
- 反馈：用户对推荐结果的交互行为（例如：点击、购买、观看时长、忽略、差评）。系统根据反馈更新用户画像和推荐模型。

#### 4. 人脸识别门禁系统

- 输入：摄像头捕捉到的人脸图像。
- 输出：识别结果（例如：匹配成功并开门、匹配失败并提示或报警）。
- 反馈：管理员对错误识别的修正（例如：将错误识别为 A 的人脸手动标记为 B）、新录入授权人员的人脸信息。系统根据反馈优化识别模型。

#### 5. 机器翻译系统

- 输入：源语言文本（例如：一段英文句子）。
- 输出：目标语言文本（例如：翻译后的中文句子）。
- 反馈：用户对翻译结果的评价或修正（例如：提供更准确的译文）、双语语料库的更新。系统根据反馈改进翻译模型。

## 1.2 人工智能代表性技术选讲

### 理论：用符号方法表示方案 5。井字棋：方案 5（课后作业）

为形式化表达 方在第  $n+1$  步的决策策略，设井字棋棋盘为  $3 \times 3$  矩阵，记为  $TTT_n$ ，其中  $TTT_n \text{Cell}_{i,j} \in -1, 0, 1$  表示第  $n$  步时  $(i,j)$  格的状态：-1 为对方 ( $\times$ )，1 为己方 ( $\circ$ )，0 为空。

**判断  $(i_1, j_1), (i_2, j_2), (i_3, j_3)$  是否为一条线的逻辑条件：**

$$\text{IsLine}(i_1, j_1; i_2, j_2; i_3, j_3) := \begin{cases} (i_1 = i_2 = i_3) & \text{(同一行)} \\ \vee (j_1 = j_2 = j_3) & \text{(同一列)} \\ \vee (i_1 - j_1 = i_2 - j_2 = i_3 - j_3) & \text{(主对角线)} \\ \vee (i_1 + j_1 = i_2 + j_2 = i_3 + j_3 = 2) & \text{(副对角线)} \end{cases}$$

本策略的三条原则如下，按优先级依次考虑：

1. **能冲三就冲三**：符号表示（假设棋盘状态为  $T^n$ ，1 表示我方，-1 表示对方，0 表示空格）：存在一条可构成三连的线  $L_i = \{(i_1, j_1), (i_2, j_2), (i_3, j_3)\}$ ，满足已有两个己方棋子，且第三格为空：

$$\exists i [TTT_n \text{Cell}_{i_1, j_1} = 1 \wedge TTT_n \text{Cell}_{i_2, j_2} = 1 \wedge TTT_n \text{Cell}_{i_3, j_3} = 0]$$

则：

$$\text{TTT}_{n+1}\text{Cell}_{i_3,j_3} := 1$$

若存在多个满足条件的位置，任选其一即可。

## 2. 防住对方下一步可能的冲三：

存在一条对方可在下一步构成三连的线  $L_i = \{(i_1, j_1), (i_2, j_2), (i_3, j_3)\}$ ，满足已有两个对方棋子，且第三格为空：

$$\exists i [\text{TTT}_n \text{Cell}_{i_1,j_1} = -1 \wedge \text{TTT}_n \text{Cell}_{i_2,j_2} = -1 \wedge \text{TTT}_n \text{Cell}_{i_3,j_3} = 0]$$

则：

$$\text{TTT}_{n+1}\text{Cell}_{i_3,j_3} := 1$$

（若规则 1 无满足项，优先阻止对方）

## 3. 随机落子：

若上述两条规则均不满足，则从空格中任选一个落子：

$$\text{Rand}(i, j) \text{ such that } \text{TTT}_n \text{Cell}_{i,j} = 0, \quad \text{set } \text{TTT}_{n+1}\text{Cell}_{i,j} := 1$$

**理论：估算围棋的搜索空间。**

围棋棋盘大小为  $19 \times 19 = 361$  个交叉点。每个点有三种可能状态：黑子、白子、无子。因此，状态空间的上界约为  $3^{361}$ 。

实际上，根据研究，约为  $2.08 \times 10^{170}$

## 1.3 人工智能简史

**理论：证明两层单调激活函数神经网络不能表达异或函数。**

**证明两层神经网络（单层感知机）不能表达异或函数：**

异或（XOR）函数的真值表如下：

$x_1$	$x_2$	$y$
0	0	0
0	1	1
1	0	1
1	1	0

一个两层神经网络（即输入层直接连接到输出层，中间没有隐藏层，或称为单层感知机）的输出可以表示为： $y = f(w_1x_1 + w_2x_2 + b)$  其中  $w_1, w_2$  是权重， $b$  是偏置， $f$  是激活函数（通常是阶跃函数或 Sigmoid 函数）。

这个公式  $w_1x_1 + w_2x_2 + b = 0$  在二维平面上定义了一条直线。这条直线将平面分为两个区域。单层感知机只能对线性可分的数据进行分类。

我们将 XOR 的四个输入点  $(0,0), (0,1), (1,0), (1,1)$  在二维平面上表示出来。点  $(0,0)$  和  $(1,1)$  对应输出 0，点  $(0,1)$  和  $(1,0)$  对应输出 1。

我们无法找到一条直线，能将  $(0,0)$  和  $(1,1)$  分在一侧，同时将  $(0,1)$  和  $(1,0)$  分在另一侧。因此，XOR 函数是线性不可分的，两层神经网络无法表达异或函数。

### 理论：假设翻译的标注一句成本 1 毛，那么把维基百科标注成中文要多少钱？

这是一个估算问题，需要对维基百科的句子数量进行估算。

1. **估算英文维基百科的文章数量：**根据维基百科统计数据（截至 2023 年底或 2024 年初），英文维基百科大约有 670 万篇以上的文章。我们取一个整数方便计算，例如  $N_{articles} = 6.7 \times 10^6$  篇。
2. **估算平均每篇文章的句子数量：**这是一个变化较大的值。短条目可能只有几句话，长条目可能有成百上千句。我们可以假设一个平均值。例如，一篇中等长度的文章可能有 50-200 个句子。我们取一个中间偏保守的值，例如  $S_{avg\_per\_article} = 100$  句/篇。（这个估算可能偏低，因为很多文章非常长，但也有很多短的存根页面）
3. **估算总句子数量：**总句子数  $S_{total} = N_{articles} \times S_{avg\_per\_article} = (6.7 \times 10^6) \times 100 = 6.7 \times 10^8$  句。即约 6.7 亿句话。
4. **计算总成本：**每句成本  $C_{per\_sentence} = 0.1$  元/句。总成本  $C_{total} = S_{total} \times C_{per\_sentence} = (6.7 \times 10^8) \times 0.1 = 6.7 \times 10^7$  元。即 6700 万元人民币。

因此，基于上述估算，将英文维基百科标注成中文大约需要 **6700 万元人民币**。这是一个非常粗略的量级估算。