

# 全国计算机等级考试（NCRE）

## 一级人工智能与大模型基础

### 样题及参考答案

#### ➤ 样题

#### 一、单项选择题

1、人工智能的提出者约翰·麦卡锡认为，人工智能的核心目标是

- A) 模拟人类智能行为并扩展其能力
- B) 开发具有自我意识的机器
- C) 完全替代人类劳动
- D) 构建纯数学推理系统

2、弱人工智能与强人工智能的核心区别在于

- A) 是否具备自主意识与决策能力
- B) 是否依赖大数据
- C) 是否使用神经网络
- D) 是否使用机器学习

3、现在人工智能已经成为一项通用技术，俗称通用人工智能，这是因为人工智能

- A) 能识别自然语言并生成固定回复
- B) 能在预设的特定领域内执行任务
- C) 依赖人工干预完成复杂任务
- D) 具备跨领域应用能力，可推动多行业技术革新

4、“深度学习革命”的关键突破是

- A) 2012 年 AlexNet 在 ImageNet 竞赛中的优异表现
- B) 1997 年深蓝计算机的发明
- C) 2016 年 AlphaGo 算法的实现
- D) 2020 年 GPT-3 的发布

5、对于当前人工智能发展阶段描述最符合的是

- A) 多模态大模型与通用人工智能探索
- B) 符号主义主导
- C) 联结主义主导
- D) 依赖人工标注数据

6、卷积神经网络 CNN 最适合应用的任务是

- A) 股票价格预测
- B) 自然语言翻译
- C) 图像分类
- D) 机器人路径规划

7、自动驾驶汽车最关键的技术挑战是

- A) 导航地图的精确性
- B) 燃油发动机的精准控制
- C) 传感器的准确感知
- D) 高精度地图与传感器融合

8、以下属于监督学习的应用场景是

- A) 有标注数据集的图像分类
- B) 基于无标签数据的聚类分析
- C) 强化学习的游戏 AI 训练
- D) 基于生成对抗网络的图像生成

9、某医院计划开发智能医疗辅助系统，利用知识图谱技术提升诊疗效率。下列功能最能体现知识图谱核心优势的是

- A) 根据患者主诉症状自动推导出可能的病因，并提供相关诊断依据和鉴别要点
- B) 自动识别医学文献中的关键词，统计疾病与症状的共现频率
- C) 存储海量患者的电子病历数据，并支持按姓名、病历号等字段快速检索
- D) 将医生的语音问诊记录实时转写为结构化文本病历

10、某在线教育平台希望利用人工智能技术提升学习体验。以下应用场景中最能体现生成式人工智能“从数据中创造新内容”的核心能力的是

- A) 为不同学习进度的学生，动态生成个性化的练习题和答案解析
- B) 根据学生的答题记录，自动判断其知识薄弱点并归类
- C) 根据预设的课程大纲和时间表，自动推送学习任务和提醒
- D) 识别学生论坛讨论中的高频问题并生成统计报告

11、以下不属于大语言模型在金融领域应用的是

- A) 信贷风险评估
- B) 股市风险预测
- C) 客户服务机器人
- D) 实时高频交易处理

12、以下不属于大语言模型核心特征的是

- A) 基于 Transformer 等深度架构的多层神经网络
- B) 无需大量训练数据
- C) 参数规模庞大
- D) 具备多任务泛化能力

13、“思维链（Chain of Thought）”技术的主要作用是

- A) 分解复杂任务为可执行步骤
- B) 提升模型训练速度
- C) 减少模型参数量
- D) 增强模型视觉识别能力

14、2025 年推出的 DeepSeek R1 的最重要突破是

- A) 推理的算力成本降低
- B) 支持多模态生成
- C) 参数规模达万亿级
- D) 实现通用人工智能

15、以下可以缓解大语言模型“偏见”的技术是

- A) 提高模型复杂度
- B) 增加训练数据量
- C) 增加模型公平性评估
- D) 减少训练轮次

16、以下不属于智能计算核心加速组件的是

- A) 网络处理器
- B) 图形处理器
- C) 神经网络处理器
- D) 张量处理器

17、某人工智能实验室需部署一套算力集群，用于训练参数量为 100 亿的 Transformer 模型。已知训练该模型需达到 1 PFLOPS (1000 TFLOPS) 的持续双精度浮点算力。可选配置如下表所示，以下方案中满足算力需求且总成本最低的是

显卡型号	单卡双精度算力	单卡价格	单机最大卡数
P100	5 TFLOPS	10,000	8
V100	15 TFLOPS	40,000	8
A100	20 TFLOPS	70,000	8

- A) 7 台服务器，每台配 8 张 A100
- B) 50 台服务器，每台配 8 张 P100
- C) 9 台服务器，每台配 8 张 V100
- D) 10 台服务器，每台配 8 张 P100

18、关于神经网络处理器 NPU，以下描述不正确的是

- A) 可以高效执行深度学习任务
- B) 主要用于处理神经网络算法
- C) 具有强大的并行计算能力
- D) 适用于所有计算任务

19、一种专为并行计算设计的高性能处理器，最初用于渲染图形，现已广泛应用于深度学习、科学计算、数据可视化等领域。符合这个描述的处理器是

- A) GPU
- B) ARM
- C) CPU
- D) NPU

20、以下不属于算力构成要素的是

- A) 数据加密能力
- B) 处理器性能
- C) 数据传输能力
- D) 并行计算能力

21、若云计算的集中处理方式无法满足计算的低时延、低功耗需求，最适合满足该需求的技术是

- A) 边缘计算
- B) 内存扩容
- C) 物联网技术
- D) 并行计算

22、在一个三分类问题中，某样本的真实标签为第一类，即 one-hot 编码为[1,0,0]，模型预测的概率分布为[0.5, 0.25, 0.25]，则交叉熵损失（以 2 为底）为

- A) 1.5
- B) 0.5
- C) 1.0
- D) 2.0

23、采用表格表达的数据属于

- A) 非结构化数据
- B) 半结构化数据
- C) 结构化数据
- D) 时序数据

24、在数据预处理中，“数据清洗”的主要目的是

- A) 去除噪声、填补缺失值、修正错误数据
- B) 将数据转换为数值型格式
- C) 对数据进行分类标注
- D) 提高数据维度以增强模型性能

25、以下不能直接提升自动化标注水平的是

- A) 数据可视化
- B) 数据质量
- C) 计算资源
- D) 算法能力

26、“数据标注”的核心作用是

- A) 减少模型训练所需的数据量
- B) 将数据转换为模型可处理的格式
- C) 为监督学习提供带标签的训练数据
- D) 为数据添加注释

27、对于 XML 数据格式的描述正确的是

- A) XML 基于标签设计
- B) XML 不支持注释
- C) XML 基于键值设计
- D) XML 不能自定义标签

28、以下不能支撑知识图谱“可解释性”的特征是

- A) 逻辑可溯
- B) 结构透明
- C) 可视化
- D) 语义赋能

29、知识表示的最核心目标是

- A) 增强数据可视化效果
- B) 提高数据存储效率
- C) 提高数据计算效率
- D) 将知识转化为机器可处理的形式

30、知识库与知识图谱的主要区别在于

- A) 知识库侧重存储，知识图谱强调图结构表示和语义关联
- B) 知识库强调实体关系，知识图谱强调静态存储
- C) 知识库支持语义推理，知识图谱仅支持关键词检索
- D) 知识库以图结构组织，知识图谱以表格形式存储

31、人工智能技术在家庭生活中的典型应用是

- A) 具备语音唤醒、内容推荐功能的智能音箱
- B) 可以无线投屏的投影仪
- C) 支持 APP 远程预约的智能电饭煲
- D) 可调节温度和风速的变频空调

32、机器学习中数据集通常划分为

- A) 训练集、验证集、测试集
- B) 非结构化数据集和结构化数据集
- C) 视频数据集、文本数据集、声音数据集、图像数据集
- D) 离散型数据集和连续型数据集

33、银行分析师依据用户年龄、婚姻、收入及房产状况，决定是否发放贷款，下列算法中最适合分析师应用的是

- A) 一元线性回归算法
- B) K-Means 算法
- C) 决策树算法
- D) 多元线性回归算法

34、K 最近邻算法 KNN 中 “K” 表示

- A) K 个数据的类别标签
- B) 随机选择 K 个初始数据点
- C) 数据分类的数量 K
- D) 在预测时需参考的最近邻样本数量

35、以下不属于 K-Means 算法关键步骤的是

- A) 选择距离最近的 K 个训练样本
- B) 分配数据点到最近质心
- C) 初始化 K 个初始质心
- D) 将簇内均值更新为质心

36、以下实际应用场景中最适合抽象为分类问题是

- A) 预测未来一周的股票价格
- B) 根据用户历史行为划分兴趣群体
- C) 判断信用卡交易是否存在欺诈
- D) 分析超市购物篮中啤酒和尿布的关联性

37、在一个人工智能预测模型中，使用线性回归预测子女身高： $Y=aX_1+bX_2+c$  ( $X_1$  为母亲身高， $X_2$  为父亲身高， $Y$  为子女身高)，模型参数  $a$ 、 $b$ 、 $c$  的获取方法是

- A) 通过带有真实子女身高标注的数据集训练得到
- B) 基于遗传学经验公式设定
- C) 通过无标注的家族身高数据聚类分析得出
- D) 通过支持向量机 SVM 分类模型得到

38、让大语言模型在海量无标签数据中学习语言规律和语义关联，从而获得通用语言理解能力。这一核心技术属于

- A) 预训练
- B) 无监督数据增强

- C) 参数微调
- D) 分布式推理加速

39、大语言模型提示工程（Prompt Engineering）的主要目的是

- A) 删除模型中的冗余参数
- B) 训练模型从零开始学习新任务
- C) 减少模型对计算资源的需求
- D) 通过设计输入指令引导模型生成符合期望的输出

40、某电商公司需要开发在线客服系统，要求能自动理解客户咨询内容，并主动追问细节以提供更精准的解决方案。以下最合适方案是

- A) 基于大语言模型和专用知识库的智能代理
- B) 预设固定“问答对”的自动回复系统
- C) 根据关键词检索数据库中的标准答案
- D) 基于通用大语言模型的对话机器人

41、某医疗科技公司需提升大模型在疾病诊断报告解读中的专业准确性，以下最有效的技术手段是

- A) 增加分布式计算节点数量
- B) 引入专家标注的疾病诊断报告数据集
- C) 引入多语言医疗文献数据集
- D) 增加无标注的 CT 和 X 光样本

42、在训练一个房价预测模型时，发现模型在训练数据上的预测误差极低，但在测试数据上误差显著增加。这种现象通常被称为

- A) 过拟合
- B) 欠拟合
- C) 数据泄露
- D) 数据增强

43、构建智能系统时，目标是在保证模型性能前提下降低对算力需求的技术是

- A) 模型并行
- B) 模型轻量化
- C) 梯度积累
- D) 早停法

44、在某医疗诊断系统中，模型对 100 名患者进行疾病预测，已知真实患病者 20 人。模型预测结果为：正确识别患者 18 人，误诊健康人为患者 5 人，该模型的召回率（Recall）是

- A. 25%
- B. 78.3%
- C. 18%
- D. 90%

45、某智能音箱误采集了用户未授权的对话录音，正确的处理方式是

- A) 立即删除数据并关闭非必要采集功能
- B) 加密存储后用于改进语音识别模型
- C) 通知用户已采集但继续使用数据
- D) 停止采集，保留已获取的数据

46、某电商平台预测商品日销量，模型对 4 个商品的真实销量和预测结果如下表所示，该模型的均方误差（MSE）是

商品编码	真实销量（件）	预测销量（件）
商品1	320	300
商品2	480	500
商品3	420	450
商品4	380	350

- A) 2600
- B) 625
- C) 650
- D) 2500

47、在某电商平台的人工智能系统中，以下措施中最符合“数据最小化原则”的是

- A) 要求用户提供身份证号以提升推荐精度
- B) 采集用户购买历史和浏览行为数据
- C) 要求用户进行人脸识别以确认身份
- D) 采集用户家庭成员信息以分析家庭购物偏好

48、某医院部署 AI 辅助诊断系统时，需处理包含患者敏感信息的历史电子病历。以下措施最符合隐私保护要求的是

- A) 对病历中的姓名、身份证号进行加密

- B) 实施动态口令认证机制保障数据访问安全
- C) 使用数据压缩技术保存患者的 X 光片
- D) 要求患者每次就诊更换用户名和密码

49、某金融机构部署 AI 信贷审批系统时，发现模型对特定地域群体存在审批通过率差异。以下措施最能有效减少算法偏见的是

- A) 每季度对不同地域群体的审批通过率进行统计检验
- B) 锁定初始训练数据集，禁止后续模型参数更新
- C) 根据用户移动端操作频率动态调整模型响应速度
- D) 要求申请人补充提交家庭成员的信用记录

50、某国产大模型厂商宣布其 API 服务价格降至行业均价的 1/20，引发市场连锁反应。以下最直接导致该降价现象的技术突破是

- A) 新技术优化推理效率，降低推理成本
- B) 政府出台 AI 算力基建专项补贴政策
- C) 用户增长放缓导致算力过剩
- D) 模型从千亿参数缩减至百亿级

## 二、多项选择题

1、关于现代处理器的功能特性，下列描述正确的是

- A) CPU 作为计算机核心组件，承担指令解析与通用计算任务
- B) GPU 凭借大规模并行架构，在 3D 渲染任务中效率显著高于 CPU
- C) NPU 主要用于推理阶段的矩阵运算加速
- D) GPU 在图像卷积运算中比 CPU 快
- E) NPU 通过硬件级并行单元，可同时处理数千个神经网络节点计算

2、以下选项属于非结构化数据的是

- A) 文本文件
- B) 图像
- C) 录音
- D) 视频
- E) 表格数据

3、以下选项属于知识图谱构建步骤的是

- A) 数据清洗
- B) 实体识别
- C) 关系抽取
- D) 集群部署
- E) 需求分析

4、关于人工智能算法的核心特性，下列描述正确的是

- A) 逻辑回归模型通过极大似然估计进行参数估计
- B) 分类算法可能输出概率值或类别标签
- C) 一元线性回归使用梯度下降优化均方误差
- D) 数据预处理仅需处理异常值，无需处理缺失值
- E) 人工智能算法的运行效果与算力无关

5、关于人工智能典型算法的特性，下列描述正确的是

- A) 通过计算数据点与簇中心的距离动态调整聚类中心
- B) 循环神经网络（RNN）更适合处理序列数据
- C) 决策树可以直观展示决策过程
- D) 线性回归输出结果是类别的标签值
- E) 逻辑回归主要处理的是回溯问题

6、关于 TensorFlow 的描述，正确的是

- A) 基于计算图的开源深度学习框架
- B) 支持动态图与静态图混合编程模式
- C) 提供分布式训练和模型部署的完整工具链
- D) 可用于大规模数据的训练
- E) 不支持大规模神经网络

7、关于提示工程的核心特征，下列描述正确的是：

- A) 通过设计结构化提示模板引导大模型输出符合预期的结果
- B) 整合角色设定、思维链引导、多模态输入等交互技术
- C) 主要应用于自然语言处理任务，无法支持图像生成场景
- D) 通过优化提示词降低模型计算复杂度，减少参数量需求
- E) 其核心方法依赖于模型架构的底层参数调整

8、在某城市的智慧安防系统中，需要从大量高清摄像头采集的视频流中快速准确地识别出异常行为，为了提升系统的效率，以下做法正确的是

- A) 采用更高效的算法模型，在保持精度的同时降低计算量
- B) 部署边缘计算设备，在摄像头端过滤无用的背景特征
- C) 动态分配算力，在人流密集区域增加 GPU 核心数量处理复杂行为特征
- D) 采用 CPU 负责系统的控制逻辑、数据管理、矩阵运算等全部任务
- E) 采用更简单的算法模型，通过降低精度来大幅减少参数量

9、LoRA 是一种高效的大模型参数微调方法，该方法在保持模型性能的同时，极大地减少可训练参数的数量。以下关于 LoRA 技术的说法正确的是

- A) 与全参数微调相比计算成本更低
- B) 适用于算力资源受限的场景
- C) 与全参数微调相比增加了训练时长
- D) 需要重新训练整个基础模型的权重
- E) 属于提示工程的一种变体

10、下列应用场景可能引发个人隐私泄露风险的是

- A) 智能手环上传用户心率、睡眠数据至云端
- B) 智能音箱通过唤醒词误触持续录音并上传云端
- C) 游戏 APP 要求用户提供人脸信息作为唯一登录验证方式
- D) 社交媒体使用聊天记录关键词定向推送广告但未告知用户
- E) 教育平台根据学生答题记录推送同类习题

### 三、综合填空题

1、知识图谱可以分为模式层和数据层，模式层定义本体的概念关系，数据层存储具体知识的事实。

三元组的基本形式有以下两种：

实体 — 关系 — 实体

实体 — 属性 — 属性值

请采用上述三元组表示方法填写空格。

知识内容是“北京是中国的首都，北京市人口约 2000 万”。

模式层：

国家 — \_\_\_\_\_ — 城市

数据层：

中国 — \_\_\_\_\_ — \_\_\_\_\_

\_\_\_\_\_ — \_\_\_\_\_ — 约 2000 万

2、某自动驾驶公司研发车载计算平台，该平台需同时处理车载传感器数据与实时图像识别等任务，其硬件架构包含多类型计算单元。请完成以下填空

(1) 激光雷达数据融合

单帧激光雷达点云数据量：每秒处理 2000 帧，每帧包含 500000 个点

单点计算复杂度：每个点需执行 5 次浮点运算

单帧计算量：\_\_\_\_\_ 次/帧

每秒总计算量：\_\_\_\_\_ 次/秒

(2) 摄像头图像处理负载

摄像头数量：4 个

单摄像头帧率：每秒拍摄 20 张图像

单帧计算复杂度：每张图像 AI 检测需 1000 次矩阵运算

单摄像头计算量：\_\_\_\_\_ 次/秒

4 个摄像头总计算量：\_\_\_\_\_ 次/秒

(3) GPU 架构优势

在这个车载计算平台中，GPU 凭借其大量\_\_\_\_\_（填写：并行/串行）计算单元的优势，能够高效处理激光雷达点云数据和图像识别中的计算任务，显著提升算力效率。

3、某电商为了优化购物平台，期望了解“用户在商品页面停留时间的长短是否会对购买率产生影响”，收集到 8 个商品页面的平均浏览时长与购物转化率数据，如下表所示。

序号	平均浏览时长 (x, 分钟)	购物转化率 (y, %)
1	1.5	2.8
2	2.0	3.5
3	2.5	4.2
4	3.0	5.0
5	3.5	5.8
6	4.0	6.5
7	4.5	7.2
8	5.0	8.0

从表格数据来看，平均浏览时长(x)与购物转化率(y)之间近似呈现线性关系，假设线性回归模型  $y = ax + b$ ，为了确定该模型的系数 a 和 b，可采用最小二乘法，通过最小化误差平方

和来得到最佳拟合参数。

(1) 根据表格数据, 平均浏览时长的均值  $\bar{x} = \underline{\hspace{2cm}}$  分钟 (保留2位小数), 方差  $Var(x) = \underline{\hspace{2cm}}$  分钟<sup>2</sup> (保留2位小数)。

(2) 根据表格数据, 平均浏览时长(x)与购物转化率(y)的协方差为

$Cov(x,y) = \Sigma [(x_i - \bar{x})(y_i - \bar{y})]/n \approx 1.96$ 。采用最小二乘法, 计算回归系数  $a = Cov(x,y)/Var(x) \approx \underline{\hspace{2cm}}$  (保留2位小数),  $b = \bar{y} - a\bar{x} \approx \underline{\hspace{2cm}}$  (保留2位小数)。

(3) 基于上述方法得到的线性模型, 当某用户平均浏览时长为 3.8 分钟时, 模型预测的购物转化率为  $\underline{\hspace{2cm}}\%$  (保留 1 位小数)。

4、ICIO 提示工程是一种用于指导 AI 系统理解和处理复杂任务的结构化方法。ICIO 通过明确的指令 (Instruction)、背景信息 (Context)、输入数据 (Input Data) 和输出指示器 (Output Indicator) 来优化与 AI 的交互。

如果希望使用大语言模型来制定一个旅游计划, 为了使大语言模型生成如图所示的结果, 请按照 ICIO 的提示工程框架填写空格部分。

#### 每日行程时间轴

##### Day 1

- ⌚ 08:00 上海虹桥→北京南站 (高铁, 553元/人×2=1106元)
- ⌚ 12:30 入住胡同民宿 (2晚共1200元)
- ⌚ 14:00 颐和园 (门票30元/人×2=60元, 地铁4元/人×2=8元)
- ⌚ 19:00 王府井晚餐 (人均80元×2=160元)

##### Day 2

- ⌚ 08:30 故宫 (已预约门票60元/人×2=120元, 公交2元/人×2=4元)
- ⌚ 13:00 景山公园 (门票2元/人×2=4元, 步行0元)
- ⌚ 16:00 798艺术区 (免费, 地铁+公交8元/人×2=16元)
- ⌚ 20:00 三里屯网红餐厅 (人均120元×2=240元)

##### Day 3

- ⌚ 09:00 天坛 (门票15元/人×2=30元, 共享单车6元)
- ⌚ 12:00 南锣鼓巷小吃 (人均50元×2=100元)
- ⌚ 14:00 国家博物馆 (免费, 地铁4元/人×2=8元)
- ⌚ 18:00 返程高铁 (553元/人×2=1106元)

#### 费用清单

类别	明细	金额
交通	高铁往返	2212元
	市内交通 (含共享)	62元
住宿	胡同民宿2晚	1200元
餐饮	正餐+小吃	660元
门票	景点门票	226元
总计		4360元 (余640元备用)

**Instruction:** 请为我制定一个\_\_\_\_\_日北京深度游计划。

**Context:** 目标人群为 2 名大学生，预算上限\_\_\_\_\_元。

**Input Data:** 出发地点为上海，偏好景点类型为历史古迹和网红打卡地，已预约第\_\_\_\_\_日上午的故宫门票。

**Output Indicator:** 需包含至少 5 个核心景点，每日交通方式不超过 2 种，住宿类型限定为\_\_\_\_\_，最终输出为每日行程时间轴和\_\_\_\_\_。

## ➤ 参考答案

### 一、单项选择题

- |      |      |      |      |      |
|------|------|------|------|------|
| 1.A  | 2.A  | 3.D  | 4.A  | 5.A  |
| 6.C  | 7.D  | 8.A  | 9.A  | 10.A |
| 11.D | 12.B | 13.A | 14.A | 15.C |
| 16.A | 17.C | 18.D | 19.A | 20.A |
| 21.A | 22.C | 23.C | 24.A | 25.A |
| 26.C | 27.A | 28.C | 29.D | 30.A |
| 31.A | 32.A | 33.C | 34.D | 35.A |
| 36.C | 37.A | 38.A | 39.D | 40.A |
| 41.B | 42.A | 43.B | 44.D | 45.A |
| 46.C | 47.B | 48.A | 49.A | 50.A |

### 二、多项选择题

- |          |         |        |        |         |
|----------|---------|--------|--------|---------|
| 31.ABCDE | 32.ABCD | 33.ABC | 34.ABC | 35.ABC  |
| 36.ABCD  | 37.AB   | 38.ABC | 39.AB  | 40.ABCD |

### 三、综合填空题

(略)