

# 人工智能应用工程师（高级）考试大纲

## （试行版）

人工智能应用工程师是能够利用人工智能相关技术进行应用研发，并开展各类工作的从业人员统称。人工智能应用工程师未来发展大致分为三个方向，并对应了职业技能考评的三个级别：

一是高级技术人员，对应人工智能应用工程师（高级），例如算法研究人员、高级架构人员；

二是业务管理人员，对应人工智能应用工程师（中级），例如产品经理、技术经理、项目经理等；

三是应用技术人员，对应人工智能应用工程师（初级），例如数据标注人员、知识库开发人员、应用开发人员等。

人工智能应用工程师（高级）的主要职能包括：1. 负责应用机器学习、深度学习、强化学习或交互技术等人工智能技术设计、研究及实现相关业务需求；2. 实现系统框架和核心代码，解决开发过程中的技术问题；3. 开展人工智能产品的规划、架构、设计和研发。

按照以上岗位职能要求，参考人工智能应用领域的行业特点提出了如下考核内容。

### 一、考核对象

1. 计算机、软件、数学、通信、人工智能、大数据以及其他相关专业有志于进入人工智能及大数据领域的学生
2. 人工智能、大数据、互联网等相关产品厂商及服务企业的管

理人员、技术人员和业务人员

3. 行业用户或企业用户的管理人员、技术人员和业务人员

## 二、考核目标

1. 了解人工智能、图灵测试的概念、特点以及二者之间的关系，掌握强弱人工智能的区别，掌握人工智能的历史沿革、各国发展战略，以及人工智能在基础层、技术层和应用层的产业发展概况
2. 掌握机器学习的概念和特点以及机器学习的任务流程，了解机器学习与深度学习、数据挖掘以及数据分析的关系
3. 掌握数据集、训练集、验证集、测试集、决策树和熵的概念以及特征，掌握监督式学习与非监督式学习、强化学习与非强化学习的概念以及区别
4. 了解神经网络的演变历史以及麦卡洛克-皮茨网络、增量规则、反向传播、离散型霍普菲尔德网络的概念和特点
5. 掌握数据清洗、探索性分析以及特征工程等数据与处理方法的概念和特点，了解分类、回归、去噪、转录、机器翻译、异常检测以及结构化输出等机器学习常见任务的概念和操作方法
6. 了解统计模型、误差和 MSE 的概念及特点，掌握迭代法和梯度下降法的概念及操作原理
7. 掌握机器学习模型的评估方法与改进策略，了解 Scikit-Learn、Spark MLlib 的概念及基本操作规范
8. 掌握算法的概念及分类以及线性回归、逻辑回归的概念及原理
9. 掌握线性判别分析的概念及特点，分类与回归树分析的概念及区

别

10. 掌握贝叶斯及朴素贝叶斯的概念及模型原理
11. 掌握 Kzui 近邻算法、矢量量化的概念和特点, 并了解 Bagging 和随机森林、Boosting 和 AdaBoost 的概念、特点及区别
12. 掌握支持向量机、逻辑回归算法、KNN 算法、集成算法和聚类算法的概念、特点及操作原理, 了解掌握算法评价原则及方法
13. 掌握神经网络、神经网络参数的概念、特点、分类以及训练神经网络的方法和优化神经网络的方法
14. 掌握深度学习、深度神经网络、正向传播的概念和特点, 同时掌握深度学习的优势
15. 了解卷积神经网络、经典卷积神经网络以及深度强化学习的概念、分类及操作原理
16. 掌握知识图谱、开放知识图谱、知识图谱融合的概念、特点、分类和架构, 并了解知识图谱在行业数据分析中的应用, 掌握知识提取、分析以及知识图谱构建的方法
17. 掌握数据挖掘的概念和特点, 了解数据挖掘技术产生的时代背景, 掌握数据挖掘与数据分析的区别
18. 了解分类、聚类、回归分析、关联规则、神经网络、Web 数据挖掘、特征分析法和偏差分析法的概念并熟悉其操作方法
19. 了解信度与效度思维的概念与区别, 了解分类思维、漏斗思维、逻辑树思维、时间序列思维、指数化思维以及循环/闭环思维的概念及特点

20. 掌握语言处理的概念和特点，了解自然语言处理的历史发展
21. 掌握转换语法、系统语法、格语法和语义语法的概念和特点，并了解统计解析、机器翻译以及词义消歧的概念及操作方法
22. 掌握统计模型方法的概念和分类，了解最大熵模型、条件随机场模型、宾夕法尼亚书库项目、Wordnet、隐喻模型的概念及原理，了解隐马尔科夫模型与条件马尔可夫模型的概念及区别
23. 掌握语义搜索、知识问答的概念和特点，掌握结构化查询语言的分类，了解语义搜索的交互范式，了解知识问答与信息检索和数据库查询的区别，
24. 掌握 TensorFlow 的概念、特点及工具包的主要内容，了解装载数据、探索数据、训练模型、评估模型、优化模型、合成特征、离群值处理的操作方法
25. 掌握计算语言的概念以及 Python 与自然语言的关系，了解获得文本获取、文本分类、信息提取、文本分析以及语言数据管理的方法

### **三、考核内容**

#### **(一) 人工智能概述**

##### **1. 人工智能的概念**

- 人工智能的概念
- 思维、智能的概念
- 图灵测试的概念，图灵测试的由来以及其与人工智能的关系
- 强人工智能与弱人工智能的区别

## 2. 人工智能的历史沿革

- 人工智能的早期历史
- 人工智能发展的里程碑：博弈、专家系统、神经计算、进化计算、自然语言处理、生物信息学的概念
- 新千年人工智能的发展特点
- 美国、英国、日本、欧盟等国家的人工智能发展战略
- 我国人工智能发展战略

## (二) 人工智能产业发展概况

### 1. 基础层

- 芯片产业发展概况（GPU、FPGA、ASIC、TPU）
- 传感器产业发展概况

### 2. 技术层

- 决策管理技术的概念及特点
- 深度学习技术的概念及特点
- 生物特征识别技术的概念及特点
- 机器人技术的概念及特点
- 自然语言处理技术的概念及特点

### 3. 应用层

- 人工智能与安防
- 人工智能与金融
- 人工智能与制造业
- 人工智能与家居

- 人工智能与医疗
- 人工智能与自动驾驶

### **(三) 机器学习**

#### **1. 机器学习概述**

- 机器学习的概念
- 机器学习的特点
- 机器学习系统中反馈的作用
- 机器学习的任务流程
- 机器学习、深度学习与人工智能的关系
- 机器学习、数据挖掘与数据分析的关系
- 数据集、特征和标签的概念
- 监督式学习和非监督式学习的概念及区别
- 强化学习和非强化学习的概念及区别
- 训练集、验证集和测试集的概念及特征
- 决策树的概念及特点
- 熵的概念及特点

#### **2. 神经网络**

- 神经网络的演变历史
- 麦卡洛克-皮茨网络的概念和特嗲
- 感知器学习规则
- 增量规则的概念
- 反向传播的概念和特点

- 实现关注点的方法
- 离散型霍普菲尔德网络的概念和特点

### 3. 数据预处理

- 探索性分析的概念和方法
- 数据清洗的方法
- 特征工程的概念和特点

### 4. 机器学习的常见任务

- 分类的概念及操作方法
- 回归的概念及操作方法
- 去噪的概念及操作方法
- 转录的概念及操作方法
- 机器翻译的概念及操作方法
- 异常检测的概念及操作方法
- 结构化输出的概念及操作方法

## (四) 统计模型

### 1. 模型概述

- 模型的概念
- 误差和 MSE 的概念及特点

### 2. 模型的训练方法

- 模型与算法的区别
- 迭代法的概念及特点

### 3. 梯度下降法

➤ 步长的概念及优化步长的方法

➤ 梯度下降法的分类

➤ 梯度下降法的算法

#### 4. 模型的拟合效果

➤ 欠拟合与过拟合的概念与特点

➤ 过拟合的处理方法

#### 5. 模型的评估与改进方法

➤ 机器学习模型的评估方法

➤ 机器学习算法的改进策略

#### 6. 机器学习的实现框架

➤ Scikit-Learn 的概念及基本操作规范

➤ Spark MLlib 的概念及基本操作规范

### (五) 机器学习算法

#### 1. 算法概述

➤ 算法的概念及特点

➤ 算法的分类

➤ 线性回归的概念及原理

➤ 逻辑回归的概念及原理

➤ 线性判别分析的概念及方法

➤ 分类与回归树分析的概念及区别

➤ 贝叶斯及朴素贝叶斯的模型原理

➤ Kzui 近邻算法的概念及原理

- 学习矢量量化的概念及特点
- Bagging 和随机森林的概念及特点
- Boosting 和 AdaBoost 的概念及特点

## 2. 算法原理

- 支持向量机的概念、特点及操作原理
- 逻辑回归算法的概念、特点及操作原理
- KNN 算法的概念、特点及操作原理
- 集成算法的概念、特点及操作原理
- 聚类算法的概念、特点、不同分类及操作原理
- 算法评价原则及方法

## (六) 神经网络

### 1. 神经网络概述

- 神经网络的概念
- 神经网络的分类
- 训练神经网络的方法：正向传播算法、反向传播算法
- 标准化特征值的概念
- 丢弃正则化的概念
- 多类别神经网络的概念及操作方法
- 神经网络参数的概念及特点
- 神经网络的优化和改进方法

### 2. 深度学习

- 深度学习的概念

- 深度神经网络的概念
- 正向传播的概念
- 激活函数的概念和分类
- 深度学习的优势

### 3. 深度前馈神经网络

- 网络的前馈方式
- 神经元与全连接结构的概念
- 激活函数的分类
- 去线性化的方法
- 多层网络解决异或运算的方法
- 损失函数的概念和定义方法

### 4. 优化网络的方法

- 基于梯度下降的优化算法
- 反向传播算法
- 指数衰减的学习率算法
- 过拟合和欠拟合算法

### 5. 卷积神经网络

- 卷积运算的概念
- 卷积层的概念及特点
- 池化的过程及常用的池化函数
- 卷积神经网络的一般框架

### 6. 经典卷积神经网络

- LetNet-5 卷积网络模型的结构和实现方法
- AlexNet 卷积网络模型的结构和实现方法
- VGGNet 卷积网络模型的结构和实现方法
- InceptionNet-V3 卷积网络模型的结构和实现方法
- ResNet 卷积网络模型的结构和实现方法

## 7. 深度强化学习

- 深度强化学习的概念
- Q 学习的概念及特点
- 深度 Q 学习的概念及特点
- 深度 Q 网络的概念及特点

### (七) 知识图谱

#### 1. 知识图谱概述

- 知识图谱的概念
- 知识图谱的架构
- 开放知识图谱的概念和特点
- 知识图谱在行业数据分析中的作用

#### 2. 知识图谱构建的关键技术

- 知识提取的概念和应用方法
- 语义提取概念和应用方法
- 属性和属性值提取的概念和应用方法
- 关系抽取的概念和应用方法
- 知识表示的概念和应用方法

### 3. 知识图谱融合

- 知识图谱融合的概念和特点
- 知识图谱中的异构分类
- 本体概念层的融合方法与技术
- 实例层的融合与匹配方法

### 4. 知识图谱推理

- 知识图谱推理概念及特点
- 基于演绎的知识图谱推理方法
- 基于归纳的知识图谱推理方法

### 5. 企业知识图谱的构建方法

#### (八) 数据挖掘

##### 1. 数据挖掘概述

- 数据挖掘的概念
- 数据挖掘技术产生的时代背景
- 数据挖掘与数据分析的区别

##### 2. 数据挖掘技术

- 分类的概念及操作方法
- 聚类的概念及操作方法
- 回归分析的概念及操作方法
- 关联规则的概念及操作方法
- 神经网络的操作方法
- Web 数据挖掘的操作方法

➤ 特征分析法的概念及操作方法

➤ 偏差分析法的概念及操作方法

### 3. 大数据思维

➤ 信度与效度思维的概念及区别

➤ 分类思维的概念及特点

➤ 漏斗思维的概念及特点

➤ 逻辑树思维的概念及特点

➤ 时间序列思维的概念及特点

➤ 指数化思维的概念及特点

➤ 循环/闭环思维的概念及特点

## (九) 自然语言处理

### 1. 自然语言处理概述

➤ 自然语言处理的概念

➤ 自然语言处理的历史发展

### 2. 句法和形式语法

➤ 语法的类型

➤ CYK 算法概述

### 3. 语义分析和扩展语法

➤ 转换语法的概念和特点

➤ 系统语法的概念和特点

➤ 格语法的概念和特点

➤ 语义语法的概念和特点

#### 4. 自然语言处理中的统计方法

- 统计解析的概念及操作方法
- 机器翻译的概念和特点
- 词义消歧的概念及操作方法

#### 5. 自然语言处理的统计模型

- 统计模型方法的概念及分类
- 隐马尔可夫模型和条件马尔可夫模型的概念及区别
- 最大熵模型的概念及原理
- 条件随机场模型的概念及原理

#### 6. 统计自然语言处理语言数据集

- 宾夕法尼亚树库项目概述及原理
- Wordnet 的概念及原理
- 隐喻模型的概念及原理

#### 7. 语义搜索

- 语义搜索的概念和特点
- 结构化的查询语言分类
- 语义搜索的交互范式

#### 8. 知识问答

- 知识问答的概念和基本要素
- 知识问答与信息检索和数据库查询的区别
- 知识问答的分类体系
- NLIDB、IRQA、KBQA、CommunityQA 等知识问答系统的概念及特点

- 知识问答的评价方法

## **(十) TensorFlow**

### **1. TensorFlow 概述**

- TensorFlow 的概念及特点
- TensorFlow 工具包的主要内容

### **2. TensorFlow 程序的主要内容**

- 装载数据的操作方法
- 探索数据的操作方法
- 训练模型的操作方法
- 评估模型的操作方法
- 优化模型的操作方法
- 合成特征的操作方法
- 离群值处理的操作方法

### **3. TensorFlow 编程策略**

- 计算图的概念与特点
- 张量的概念和特点
- 会话的概念和特点
- 使用 TensorFlow 创建变量的方法
- 变量与张量的区别

## **(十一) Python**

### **1. Python 概述**

- Python 的概念与操作规范

- Python 与自然语言处理
- 计算语言的概念
- 2. 文本获取**
- 获得文本语料和词汇资源
- 原始文本的处理方法
- 编写结构化程序的方法
- 3. 文本分类及信息提取**
- 分类和标注词汇
- 学习分类文本
- 从文本提取信息
- 4. 文本分析**
- 分析句子结构
- 建立基于特征的文法
- 分析语句的含义
- 5. 语言数据管理**
- 语料库的结构
- 语料库的生命周期
- 语料库的卡发方法
- XML 的使用方法
- Toolbox 数据的特点及使用方法

\*具体考核内容以最终官方发布为准