

## 一、信息技术基础

### (一) 数据、信息与知识

#### 1. 数据

- 定义：对客观事物的符号表示（如数字、字符、图像、音频、视频等）。
- 示例：温度传感器记录的数值、学生考试成绩表格。

#### 2. 信息

- 定义：经过加工处理后的有意义的数据。
- 特征：
  - 载体依附性（需依附于介质，如纸张、硬盘）。
  - 价值性（信息对决策的作用）。
  - 时效性（信息随时间推移可能失效）。
  - 共享性（可被多人同时使用）。
  - 真伪性（需验证信息的真实性）。

#### 3. 知识

- 定义：对信息进行系统化组织后形成的指导实践的认识体系。
- 转化关系：
  - 数据 → 信息（如原始温度数据 → 天气趋势分析）。
  - 信息 → 知识（如天气趋势 → 制定农业防灾策略）。

#### 4. DIKW 模型

- 数据 (Data) → 信息 (Information) → 知识 (Knowledge) → 智慧 (Wisdom) 的递进关系。

### (二) 文本、图像和声音数据的编码原理及存储容量计算

#### 1. 文本编码

- **ASCII 码**：7 位编码，表示 128 个字符（英文、符号）。
- **UTF-8**：可变长编码（1~4 字节），支持多语言。
- **存储容量计算**：

存储容量 (字节) = 字符数 × 编码位数 / 8

- 示例: 500 个汉字 (UTF-8 编码, 每个汉字 3 字节) →  $500 \times 3 = 1500$  字节。

## 2. 图像编码

- **位图 (BMP)**: 记录每个像素的颜色值, 文件体积大。
- **矢量图**: 通过数学公式描述图形, 可无限缩放不失真。
- **存储容量计算**:

存储容量 (字节) = 分辨率 (宽 × 高) × 颜色深度 (位) / 8

- 示例: 1024 × 768 分辨率, 24 位颜色深度 →  $1024 \times 768 \times 24 \div 8 = 2,359,296$  字节 ≈ 2.25MB。

## 3. 声音编码 (PCM)

- **参数**:
  - 采样频率 (如 44.1kHz)。
  - 量化位数 (如 16 位)。
  - 声道数 (单声道/立体声)。
- **存储容量计算**:

存储容量 (字节) = 采样频率 × 量化位数 × 声道数 × 时间 (秒) / 8

- 示例: 44.1kHz 采样率, 16 位量化, 双声道, 1 分钟音频 →  $44100 \times 16 \times 2 \times 60 \div 8 = 10,584,000$  字节 ≈ 10.1MB。

## (三) 大数据相关概念、处理及可视化表达

### 1. 5V 特性

- **Volume (大量)**: 数据规模庞大 (TB 级至 PB 级)。
- **Velocity (高速)**: 数据生成与处理速度快 (如实时交易数据)。
- **Variety (多样)**: 数据形式多样 (结构化、半结构化、非结构化)。
- **Value (低价值密度)**: 需挖掘有效信息 (如监控视频中提取异常行为)。
- **Veracity (真实性)**: 数据质量与可信度 (如虚假评论过滤)。

### 2. 处理流程

- **采集**: 从传感器、日志、数据库等获取数据。
- **存储**: 使用分布式存储 (如 Hadoop HDFS)。
- **清洗**: 去除重复、缺失、错误数据。
- **分析**: 运用统计分析、机器学习算法。
- **可视化**: 通过图表 (柱状图、热力图) 展示结果。

### 3. 工具应用

- **Excel**: 数据透视表 (分类汇总)、排序/筛选 (提取特定数据)。
- **Python 库**: Pandas (数据清洗)、Matplotlib (可视化)。

## (四) 人工智能关键技术及应用

### 1. 核心技术

- **机器学习**:
  - 监督学习 (标签数据训练, 如房价预测)。
  - 无监督学习 (聚类分析, 如客户分群)。
  - 强化学习 (动态决策, 如 AlphaGo)。
- **深度学习**:
  - 卷积神经网络 (CNN, 图像识别)。
  - 循环神经网络 (RNN, 自然语言处理)。

### 2. 应用场景

- **医疗**: AI 辅助诊断 (如肺部 CT 影像分析)。
- **交通**: 自动驾驶 (路径规划、障碍物识别)。
- **金融**: 风险评估 (信用评分模型)。

### 3. 伦理与挑战

- **算法偏见**: 训练数据偏差导致结果歧视 (如人脸识别对特定种族准确率)。
- **隐私保护**: 数据匿名化、联邦学习技术。

## (五) 信息社会的基本内涵、特征、应用及影响

### 1. 内涵

- 以信息技术为核心，信息产业为主导（如互联网经济）。

## 2. 特征

- **数字化**：信息以二进制形式存储（如电子档案）。
- **网络化**：全球互联（如物联网设备通信）。
- **智能化**：AI 驱动的自动化（如智能客服）。

## 3. 应用

- **经济**：电子商务（如淘宝、京东）。
- **政治**：电子政务（在线政务服务）。
- **教育**：远程教育（MOOC 平台）。

## 4. 挑战

- **数字鸿沟**：城乡/代际技术资源差距。
- **信息过载**：筛选有效信息的难度增加。

# 二、算法与 Python 程序实现

## （一）Python 语言基础

### 1. 数据类型

- **基本类型**：
  - int（整数）、float（浮点数）、str（字符串）。
  - list（列表，可修改）、tuple（元组，不可修改）、dict（字典，键值对）。
- **类型转换**：
  - `int("123") → 123`, `str(456) → "456"`。

### 2. 运算符

- **算术运算符**：+、-、\*、/、//（整除）、%（取余）、\*\*（幂）。
- **逻辑运算符**：and、or、not。

### 3. 输入输出

- `input()`：接收用户输入（默认字符串类型）。
- `print()`：格式化输出（如 `print(f"结果: {x:.2f}")`）。

## (二) 控制语句

### 1. 条件语句

```
if score >= 90:  
    print("优秀")  
  
elif score >= 80:  
    print("良好")  
  
else:  
    print("需努力")
```

### 2. 循环语句

- **for 循环**: 遍历列表、字符串等。

```
for i in range(5): # 输出 0~4  
    print(i)
```

- **while 循环**: 条件控制循环。

```
sum = 0  
  
i = 1  
  
while i <= 100:  
    sum += i  
    i += 1
```

## (三) 函数与算法

### 1. 函数定义

```
def add(a, b):  
    return a + b
```

### 2. 递归算法

- **阶乘计算**:

```
def factorial(n):  
    if n == 1:  
        return 1
```

else:

```
    return n * factorial(n-1)
```

### 3. 排序算法

- 冒泡排序:

```
def bubble_sort(arr):
```

```
    for i in range(len(arr)):
```

```
        for j in range(len(arr)-i-1):
```

```
            if arr[j] > arr[j+1]:
```

```
                arr[j], arr[j+1] = arr[j+1], arr[j]
```

## 三、信息系统组成与应用和信息安全

### (一) 网络技术基础

#### 1. TCP/IP 协议簇

- IP 协议: 负责网络层寻址与路由。
- TCP 协议: 可靠传输 (三次握手、流量控制)。
- UDP 协议: 快速传输 (无连接, 适用于视频流)。

#### 2. IP 地址与子网划分

- IPv4 分类: A 类 (1.0.0.0~126.255.255.255)、B 类 (128.0.0.0~191.255.255.255)。
- 子网掩码: 255.255.255.0 表示前 24 位为网络号。

### (二) 网络安全

#### 1. 常见威胁

- DDoS 攻击: 通过海量请求耗尽目标资源。
- 钓鱼攻击: 伪造网站窃取用户信息。

#### 2. 防护措施

- 防火墙: 过滤非法访问。
- 加密技术: HTTPS (SSL/TLS 协议) 保护数据传输。

### (三) 组网技术

## 1. 网络拓扑

- **星型拓扑**：中心节点（如交换机）连接所有设备。
- **网状拓扑**：设备间多路径连接（高冗余）。

## 2. 路由器配置

- **端口转发**：将外部请求映射到内网服务器。
- **VLAN 划分**：隔离不同部门网络，提升安全性。