# 错题整理

## 建平：

（10）小明为了完成智慧校园研究报告，下载了一些学习资料和软件工具等，以下不能进行有损压缩的是（ ）

A.智慧校园.exe B.智慧校园宣传片.AVI C.智慧校园海报.bmp D.智慧校园项目调查问卷.docx

.exe是可执行文件，.docx是文档格式，通常这两种格式不进行有损压缩。故答案为：AD。

（11）信息系统为同学们带来便利的同时，也存在一定的安全风险，比如病毒破坏、数据丢失、隐私泄露等。针对不同的风险，以下应对措施合理的是（ ）

A.安装防火墙并定期用杀毒软件查杀 B.定期对系统的数据进行备份

C.建立网络安全制度，规范网络管理流程 D.设置访问控制，只允许管理员访问系统

安装防火墙和定期杀毒可防范病毒破坏；定期备份数据可防止数据丢失；建立网络安全制度和规范流程有助于整体的网络安全管理。只允许管理员访问系统会极大限制正常使用，不是合理的应对措施。故答案为：ABC。

（5）校友使用该系统添加返校预约时，填写并提交的预约信息数据的流向是（ ）

A.业务逻辑层→数据访问层→数据库→用户界面层 B.数据访问层→业务逻辑层→数据库→用户界面层

C.业务逻辑层→用户界面层→数据访问层→数据库 D.用户界面层→业务逻辑层→数据访问层→数据库

（5）用户填写并提交预约信息时，数据的流向是从用户界面层开始，经过业务逻辑层处理，然后到数据page number 8访问层与数据库交互。故正确答案为：选项D。

（6）小申设计了系统所需的数据库，并用数据表jperInfo存放校友返校预约信息，如下图所示，则手机号phoneNumber字段应该设置为\_\_\_\_类型

（6）手机号通常包含数字、加号等字符，因此应该使用文本类型来存储。故正确答案为：选项B。

（7）小申为了提供对数据库的数据访问，编写了自定义函数于程序db.py中，该程序文件对应系统的软件架构中的\_\_\_\_层

A.用户界面 B.数据访问 C.业务逻辑 D.网络传输

（7）自定义函数用于数据访问，因此它对应的是数据访问层。故正确答案为：选项B。

（9）小申使用Flask框架搭建网站服务端，主程序run.py中的关键代码的顺序应该是（ ）

① app= Flask(\_\_name\_\_)

② from flask import Flask, render\_template, request, redirect

③ app.run()

④ @app.route(’/’)

A.②①④③ B.②③④① C.②④①③ D.②①③④

Flask程序的关键代码顺序应该是先导入Flask和其他需要的模块，然后创建Flask应用实例，接着定义路由和视图函数，最后运行应用。故正确答案为：选项A。

（10）小申编程实现校友“添加返校预约”功能，并通过在浏览器中输入URL（http://127.0.0.1:8000/add）page number 4

进行本地测试，请帮其补全下方关键代码。

|  |
| --- |
| @app.route("①\_\_\_\_ ", methods=["GET","POST"])  def add\_info():  if request.method=="②\_\_\_\_ ":  xm= request.form["name"]  xb=request.form["gender"]  phone=request.form["phoneNumber"]  year=request.form["grad Year"]  bzr= request.form["headTeacher"]  db.add\_one(xm, xb, phone, date, year, bzr)  return redirect('/')  return render\_template("add.html")  app.run(debug=True, host=’0.0.0.0’, port=③\_\_\_\_) |

（10）①根据浏览器中输入URL（http://127.0.0.1:8000/add）可知路由函数的路径为/add。故正确答案为：/add。②检查请求方法，是否是POST方法。故正确答案为：POST（须大写）。③根据浏览器输入的地址，可知端口号为：8000。故正确答案为：8000。

（12）在系统运行与测试阶段，小梅提出——所有用户都能进入“查看所有预约”功能页面不合理。为解决该问题，进一步完善该系统，可以采用的手段有\_\_\_\_。

A.身份认证 B.访问控制 C.修补漏洞 D.定期查毒

（12）在系统运行与测试阶段，针对小梅提出的问题——所有用户都能进入“查看所有预约”功能页面不合理，为了解决该问题，需要采取的手段是确保只有授权用户能够访问特定的资源或功能。这通常通过身份认证和访问控制来实现。身份认证用于确认用户身份，而访问控制则用于基于用户身份和权限来决定用户是否能够访问某个资源或执行某个操作。因此，正确答案为：选项AB。

## 2023浦东：

（8）整理预测鸢尾花算法的步骤

①将最小值对应的鸢尾花类别判定为测试数据的分类

②将机器判定的分类与测试集中的标记分类进行对比，计算正确率

③计算每种鸢尾花的属性平均值

④计算测试数据与每种鸢尾花属性平均值之间欧式距离的平方值，找到其中的最小值。

请将正确的流程步骤填写完整：\_\_\_\_→\_\_\_\_→\_\_\_\_→\_\_\_\_

8）③计算每种鸢尾花的属性平均值：首先计算每种鸢尾花的属性平均值作为分类依据。①将最小值对应的鸢尾花类别判定为测试数据的分类：根据距离最小的属性平均值确定测试数据的分类。④计算测试数据与每种鸢尾花属性平均值之间的欧式距离的平方值：用于衡量测试数据与各类别的相似度。②将机器判定的分类与测试集中的标记分类进行对比，计算正确率：最后验证机器学习模型的准确性。这些步骤组成了一个典型的监督学习算法的预测流程，用于识别和分类鸢尾花。故答案为：③①④②。

（1）信息系统是以（1）为手段，以信息需求为目标，实现对信息有效采集和输入、传输、存储、处理、输出和利用而建立的包括人、数据、硬件、软件在内的综合系统。（1）处应填入（ ）

A.现代信息技术 B.现代通信技术 C.计算机技术 D.传感技术

现代信息技术涵盖了多种技术手段，包括通信技术、计算机技术、传感技术等，能够实现对信息的有效处理，是信息系统建立的手段。故答案为：A。

（3）小申把访问系统的网址制成二维码。如图所示，使用手机扫描二维码获取了输入界面的链接信息，这种输入方式属于（ ）page number 1



图1

A.人工输入 B.自动采集 C.自动控制 D.智能识别

（3）使用手机扫描二维码获取链接信息，属于自动采集，无需人工手动输入具体信息。故答案为：B。

（5）系统运行过程中，小申发现任何用户都可以访问该信息系统的汇总页面。请你针对这种信息安全风险提出防范措施。

你选择的是（ ）；你的具体建议是\_\_\_\_。

A.身份认证。身份认证是用户在使用信息系统时，系统对用户身份的识别和确认过程，是保证信息系统安全的基本措施。

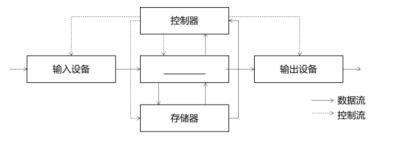
B.访问控制。访问控制是信息系统安全防范的常用技术，它的主要作用是保证信息资源不被非法访问和使用。

C.数据加密。在信息安全中，信息的保密性是指保证信息不被非授权访问，或者被非法取得数据时也无法使用该数据。数据加密就是防止数据在传输过程中被窃听，或者在存储过程中被非法访问的一种有效防范手段。

D.修补漏洞。信息系统软件存在安全风险，漏洞扫描既是攻击者探寻、入侵主机的一种最常见手段，也是管理员检测系统漏洞、发现安全隐患、提高系统安全性的有效方法。

（5）选择身份认证的原因：身份认证是信息系统安全的基本措施之一，可以有效防止未经授权的用户访问系统和敏感信息。在这种情况下，任何用户都能访问汇总页面可能导致信息泄露或非法使用，身份认证能够确认用户身份合法性，限制系统访问权限。故答案为：A。具体建议的实施意义：通过用户名和密码的身份认证，可以确保只有授权的用户才能够访问信息系统的汇总页面。这样可以有效防止未经授权的访问和潜在的安全风险，保护信息系统和数据的安全性和完整性。

（6）存储程序式体系结构的计算机工作原理如图所示，横线处的部件名称是（ ）



A.CPU B.计算器 C.运算器 D.处理器

（6）横线处的部件名称是运算器，运算器用于进行算术和逻辑运算。故答案为：C。

## 2022金山

（10）小申家种植了鸢尾花，他测量得到一株鸢尾花A的数据（花瓣长3.3厘米，花瓣宽1厘米，萼片长5厘米，萼片宽2.3厘米）。在信息技术课上，小申获得了如下表所示的三种鸢尾花属性的平均值，那么采用欧式距离的平方可以预测这株鸢尾花A为（ ）page number 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 种类 | 花瓣长 | 花瓣宽 | 萼片长 | 萼片宽 |
| 山鸢尾 | 1.4 | 0.2 | 5 | 3.4 |
| 变色鸢尾 | 4.3 | 1.3 | 6 | 2.7 |
| 维吉尼亚鸢尾 | 5 | 1.9 | 6.6 | 2.9 |

A.山鸢尾 B.变色鸢尾 C.维吉尼亚鸢尾 D.无法预测

（10）通过计算鸢尾花A与三种鸢尾花属性平均值的欧式距离的平方，比较得出与变色鸢尾的距离最小，预测为变色鸢尾。故答案为：B。

（6）小申家里安装了100M宽带，由于家中的移动设备增加，网速越来越慢。现宽带服务商提供了宽带升级体验服务，免费升级至200M，升速期3个月，到期后可续费继续使用。下列说法中错误的是（ ）

A.200M其实是200\*1024\*1024bps B.200M其实是200Mbps

C.200M其实25MB/s D.200M其实是25600Kb/s

1. A选项，200M即200Mbps，换算为字节是200÷8=25MB/s，不是200\*1024\*1024bps；B选项，200M通常指200Mbps，正确；C选项，200M换算为字节是25MB/s，正确；D选项，200Mbps换算为Kbps是200\*1024=204800Kbps，不是25600Kb/s。故答案为：AD。

（2）小申正在对运动会管理系统的数据库进行设计，建立了学生基本信息表、运动会比赛项目表、学生比赛成绩表等多个二维表，并以学生的学籍号作为关键字段建立了表和表之间的联系。这种数据库我们称之为\_\_\_\_，二维表中的每一行成为一个\_\_\_\_。

（2）小申正在对运动会管理系统的数据库进行设计，建立了学生基本信息表、运动会比赛项目表、学生比赛成绩表等多个二维表，并以学生的学籍号作为关键字段建立了表和表之间的联系。这种数据库我们称之为关系型数据库，二维表中的每一行成为一个记录。

（5）如图1所示，小申使用Python3环境中的Flask框架开发的运动会管理系统由若干模块组成。登录模块和查询模块的作用如图2所示。“查询界面”属于（ ）

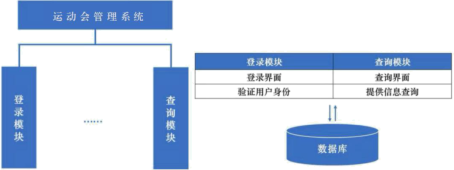


图1 图2

A.业务逻辑层 B.数据访问层 C.用户界面层

（5）业务逻辑层位于三层中的中间层，起到数据交换中承上启下的作用，用于对业务逻辑的封装；数据访问层主要负责与数据库的交互，运行数据库查询并执行更新操作，包括对表的查询、插入、更新、删除；用户界面层是软件系统与用户展示软件系统的输出信息。“查询界面”属于用户界面层。选C。

## 2023第三场

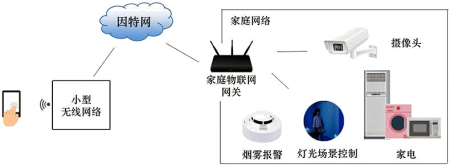
2. 如图1所示，小申家里的各类家电、摄像头及照明等物件接入智能家居系统，为日常生活带来便利。请运用所学内容，与小申一起分析以下问题。

图1

（1）小申使用手机APP远程控制摄像头查看家中的情况。从该功能的实现分析信息系统的组成要素，除软件外还有（ ）。

A.数据 B.APP C.人 D.硬件

【详解】本题考查信息系统。（1）硬件：是信息系统的基础，包括所有物理设备，如摄像头、手机、服务器等。在这个场景中，摄像头是关键的硬件，用于捕捉家中的视频数据。数据：是信息的表示和存储形式，通常被处理和解释以产生信息。在这个场景中，摄像头捕捉的视频数据是关键。人：是信息系统的用户或操作者。小申是用户，他使用APP来控制摄像头和查看视频数据。故选择ACD。

（4）图1中的“小型无线网络”连接示意图如图2所示。已知线路1使用的是双绞线，则计算机2与设备A之间的距离一般不超过\_\_\_\_\_\_\_\_\_\_\_\_米。

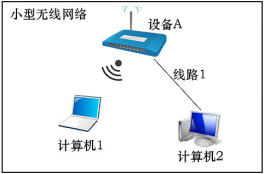
page number 2

图2

（4）双绞线组建的小型无线网络，一般不超过100米。

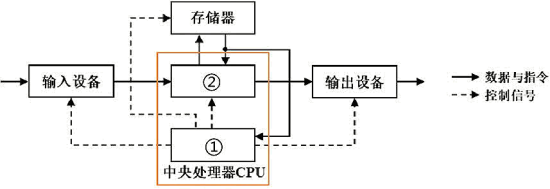
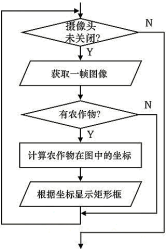
（5）计算机在信息系统中发挥着极其重要的作用。计算机由五大逻辑部件组成，如图3所示①处的逻辑部件名称为\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_。

图3

（5）根据图中控制信号虚线可知，为控制器。

## 2023第二场

（4）智能农业管理平台中的“农作物检测功能”部分算法描述如流程图所示，其使用的算法控制结构是（ ）page number 0



A.分支嵌套循环结构 B.分支嵌套分支结构 C.循环嵌套分支结构

（4）根据流程图的数据流向可知，是循环嵌套分支结构。故选择C。

（7）人工智能技术产生的巨大推动力，促使传统农业发生了巨大变化。关于人工智能对农业的影响，以下表述正确的是（ ）

A.传统农业中的一些简单重复的工种会逐渐被取代 B.需要部分从业者具有更加系统性、创造性的思维page number 1

C.智能农场使用的设备安全可靠无需担心安全问题 D.农场的管理水平与工作效率能够得到大幅度提升

（7）选项B，虽然人工智能的引入会改变农业的工作内容和需求，但说需要部分从业者具有更加系统性、创造性的思维并不完全准确。这种需求可能存在于农业技术的研发、管理和创新等高端领域，但并非所有从业者都需要这种转变。选项C，这个表述过于绝对。任何技术设备都存在一定的安全风险，包括智能农场使用的设备。虽然这些设备在设计时考虑了安全性，但用户仍需注意操作和维护，以确保其安全可靠。选项D，人工智能可以通过提供数据支持、优化决策过程和提高工作效率等方式，间接提升农场的管理水平和工作效率。然而，这并非人工智能本身的直接效果，而是其应用的结果。故选择A。

2. 如图1所示，小申家里的各类家电、摄像头及照明等物件接入智能家居系统，为日常生活带来便利。请运用所学内容，与小申一起分析以下问题。



图1

（1）从信息系统的组成要素分析智能家居系统，除硬件、软件和人以外，还有（ ）

A.网络协议 B.数据 C.通信线路 D.因特网

（1）信息系统组成要素除了硬件、软件和人以外，还有数据。故选择B。

（6）计算机在信息系统中发挥着极其重要的作用。计算机由五大逻辑部件组成，如图2所示②处的逻辑部件名称为（ ）

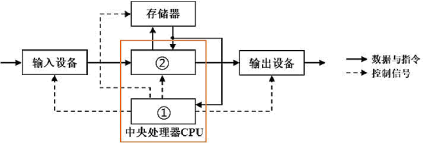


图2

A.计算器 B.运算器 C.处理器 D.控制器

（6）数据与指令的流向可知是运算器。

（4）系统中出借模块功能是登记设备借用人、使用班级和出借时间。为此小申设计了出借操作的网页界面，如图所示。以下表述正确的是（ ）



A.借用人文本框不能输入借用者真实姓名 B.出借时间的输入可以调用计算机系统的当前时间

C.设备编号和设备名称都应设计为由用户输入 D.使用班级设计成下拉列表形式是为了用户输入方便

（4）可以调用当前系统时间自动输入正确，故选择B。

## 2023第五场

（4）某次5G无线网络实验数据显示：在车辆时速100km/h的移动环境下，可以实现接近1.2Gbps的5G网络下载速度。在这样的实验环境下，从云端下载3GB的视频文件理论上大约需要\_\_\_\_\_\_秒。

（4）3GB=24Gb，24Gb/1.2Gbps=20s。故答案为20

（6）车联网系统在使用过程中会产生海量的数据，这些数据不仅产生得快，而且这些数据可能要求在几分钟内完成处理。这体现的大数据特征是（ ）。

A.数据类型多 B.价值密度低 C.处理速度快

（6）海量数据几分钟处理完，体现了大数据处理的特征处理速度快。

3. 小申所在的居委会想为银发老人建设“夕阳红”阅读驿站。作为社区志愿者，他想为阅读驿站设计一个小型的信息系统，方便老人们预约座位和查阅书籍。请运用所学内容，与小一起分析以下问题。

（1）对阅读驿站服务系统规划和分析后，小申还需要确定登录、预约、用户管理等模块以及它们之间的调用关系。其中，用户管理模块需要包含（ ）。

A.用户查询 B.新增用户 C.信息修改 D.用户审核

（1）用户管理模块应该包含用户查询、新增用户、信息修改、用户审核等功能。故本题答案是：ABCD。

## 2023第四场

（5）岩石是火星车行驶时的主要障碍之一，从相机拍摄的图像中识别出岩石是“祝融号”的一项重要“技能”。如果你是设计者，已获取了1000张原始火星岩石图像样本，请简述使用监督学习实现此识别功能的关键步骤。

答：\_\_\_\_\_\_\_\_\_\_\_\_

（5）收集、处理、分类

数据收集：第一步是收集包含输入特征和相应标签（目标变量）的训练数据集。这些标签是模型需要学习并预测的结果。数据预处理：对收集到的数据进行清洗和处理，包括处理缺失值、异常值、特征缩放、编码分类变量等，确保数据的质量并使其适合模型使用。数据集划分：将整个数据集划分为训练集和测试集（有时还包括验证集）。训练集用于训练模型，而测试集用于评估模型的性能。选择模型：根据问题的性质选择合适的监督学习模型。常见的监督学习模型包括线性回归、逻辑回归、决策树、支持向量机（SVM）、随机森林、神经网络等。模型训练：使用训练集对选定的模型进行训练。模型通过学习输入特征与标签之间的关系，调整其内部参数以最小化预测误差。模型评估：使用测试集评估模型的性能。常用的评估指标包括准确性、精确度、召回率、F1分数等。评估的目的是了解模型在未见过的数据上的表现。调整模型：根据评估结果调整模型的超参数，以提高其性能。这可能涉及到使用交叉验证、网格搜索等技术来找到最佳的超参数组合。预测新数据：当模型经过训练并调整后，可以用于对新的、未标记的数据进行预测。这是监督学习的最终目的，即利用训练好的模型对新数据进行分类或回归预测。

1. 某次5G无线网络实验数据显示：在车辆时速100km/h的移动环境下，可以实现接近1.2Gbps的5G网络下载速度。在这样的实验环境下，从云端下载2GB的视频文件理论上大约需要\_\_\_\_\_\_\_\_\_\_\_\_秒。

（4）1Gbps（gigabits per second） 等于125MB/s（兆字节每秒），下载2GB的视频文件理论上大约需要：2\*1024/1.2\*125≈13.653333，故本题答案是：13.7。

答案有问题，不能这么解，正确解法：2GB=16Gb,16/1.2=13.3s

（2）小申对阅读驿站服务系统规划和分析后，还需要确定登录、预约、系统管理等模块以及它们之间的调用关系。其中，用户座位预约模块需要包含（ ）。

A.信息查询 B.信息录入 C.信息修改 D.信息审核

（2）老人根据座位预约之前需要查询出座位的使用情况。故选择A。

（3）小申设计的阅读驿站服务系统结构示意图如图所示，包含软件的硬件设备是（ ）。



1. 无线路由器 B.便携式计算机 C.数据库 D.服务器

（3）服务器要部署软件，包含了软件的硬件。故选择D。

（4）座位预约服务功能界面如图所示。阅读驿站服务系统将会对郝爱国老人提交的这条预约信息进行（ ）。page number 5



1. 数据恢复 B.数据计算 C.数据检验 D.数据加密
2. 信息涉及到隐私信息，例如身份证号，故需要加密处理。故选择D。

## 2023 建平第一学期

（9）除了识别手写数字，人工智能技术在很多领域发挥着重要的作用。以下关于人工智能技术的应用描述正确的是（ ）

A．人工智能技术应用于智能农场、智能果园、农产品加工智能车间等，将有助于提升农业发展的智能化水平。

B．人工智能技术的应用会对一些行业和工种造成一定的影响，导致现有职业的消失，但与此同时与人工智能技术相关的岗位也在不断出现。

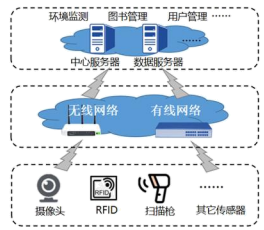
C．人工智能技术应用在面部解锁、刷脸支付、身份验证等场景，为人们的生活带来了许多便利，同时也应当确保个人私密信息的安全。

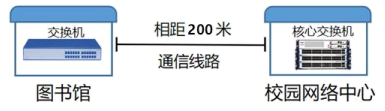
D．人工智能技术应用于自动驾驶平台，可以为用户提供覆盖广、高自动化的高精度地图服务，有助于用户更快地研发、测试和部署自动驾驶车辆。

（9）随着人工智能技术的发展，其在很多领域发挥着重要的作用。ABC选项说法正确。自动驾驶技术是一个涉及多个领域的复杂技术，人工智能技术是其中重要的一环。在自动驾驶中，人工智能主要负责实现自主决策和智能感知。其中，自主决策涉及到在各种不同的驾驶情境下，根据各种因素做出最佳的决策。这些因素包括道路情况、交通情况、天气情况、行人和其他车辆的行动，以及其他各种因素。而智能感知则主要负责实现对周围环境的感知，包括车辆和行人的位置、速度、方向等信息的获取和分析，这些信息将为自动驾驶汽车用来做出最佳的决策和行动提供支持。D选项说法错误。故本题答案是：ABC。

## 2023青浦

3. 小申学校的智能图书馆采用了物联网技术，其基本架构如图1所示。馆内各类计算机设备、图书、顶灯等物件都接入了网络，实现智能管理。请运用所学知识，与小申一起分析以下问题。



（5）如图所示，若要将图书馆内的交换机与校园网络中心的核心交换机相连接，应选用的通信线路是\_\_\_\_（双绞线/光纤），它的优势是：\_\_\_\_（3个）。

（5）常用的网络传输介质有双绞线和光纤。双绞线使用电信号传输数据，是通过电流在铜线中传输；光纤使用光信号传输数据，利用光的全反射原理，光信号在光纤内部反射前进。双绞线的传输距离有限，通常在百米以内，适合短距离内的数据传输；光纤的传输距离远，可以传输数十公里乃至数百公里，非常适合长距离数据传输。双绞线的传输速率相对较低，最高只有几百Mbps；光纤的传输速率快，可以达到数Gbps以上，适合高速数据传输。如图所示，图书馆与网络中心的距离为200米，适合用光纤作为传输介质。故⑥处答案是：光纤，⑦处答案是：1.可稳定地传输数据2.可远距离传输数据 3.可高速传输数据（合理即可）。

（9）小申家里的带宽是100Mbps，相当于每秒传输（ ）Bytes的数据。

A.100/8 B.100\*1024\*1024 C.100\*1024\*1024/8 D.100\*1024/8page number 5

1. Mbps是megabits per second的缩写，是一种传输速率单位，指每秒传输的位（比特）数量。带宽是100Mbps，相当于每秒传输100\*1024\*1024/8Bytes的数据。故本题答案是：C。

4. 小申在学习了杜甫的诗作《登高》后，想更多地了解杜甫，他和家人利用假期自驾前往位于成都的杜甫草堂参观学习。

请根据以上情境回答以下问题：

（1）小申一家到达成都后首先来到预定宾馆的智能停车场停车。在入口处，小申看到显示屏上出现了自己家车的车牌号。关于车牌号采集和显示的过程，以下说法正确的是（ ）



智能停车场入口示意图

A.小申在预定宾馆时填写了车牌号和到达时间，所以显示屏在他们到达的时间显示了他们的车牌号

B.车牌号码拍摄识别设备中有一台嵌入式计算机，可以将拍摄的车牌号码图片转换为文字

C.将车牌号码识别为文字属于人工智能中研究如何使机器“看清”和“看懂”的学科，该学科称为计算机视觉

D.车牌号码拍摄识别设备会将采集到的车牌号码、入场时间一并传输到中心服务器存储

（1）车牌识别是计算机视频图像识别技术在车辆牌照识别中的一种应用，即从图像信息中将车牌号码提取并识别出来。计算机视觉是一门研究如何使机器“看”的科学，更进一步的说，就是是指用摄影机和电脑代替人眼对目标进行识别、跟踪和测量等机器视觉，并进一步做图形处理，使电脑处理成为更适合人眼观察或传送给仪器检测的图像。车牌号码拍摄识别设备中有一台嵌入式计算机，可以将拍摄的车牌号码图片转换为文字；将车牌号码识别为文字属于人工智能中研究如何使机器“看清”和“看懂”的学科，该学科称为计算机视觉。由于车牌识别技术的应用，小申不需要提前填写车牌号和到达时间。故本题答案是：BC。

## 2024闵行

（3）小申下载了一些气象图片，某张256色的BMP图片的分辨率为1024\*768，理论上，该图片所占存储空间为\_\_\_\_\_\_MB。

（3）图像文件存储量的计算公式：存储量=水平像素×垂直像素×每个像素所需位数/8（字节）。256色的BMP图片的分辨率为1024\*768，其存储容量为：1024\*768\*8/8/1024/1024=0.75MB。故本题答案是：0.75MB。

2. 小申发现一种对黑白图像（颜色编号：黑色1，白色0）的压缩方法。压缩规则是：①记录原数据第一个位置的颜色编号：②从左往右依次扫描颜色编号，统计并记录连续出现的相同颜色的编号个数。例如图像的颜色编号为：0000000011111000，压缩结果为“0，8，5，3”。据此完成下面小题。

（1）这种压缩方法属于\_\_\_\_\_\_\_\_\_\_\_\_（选填：有损压缩/无损压缩）。

1. 有损压缩是利用了人类对图像或声波中的某些频率成分不敏感的特性，允许压缩过程中损失一定的信息：虽然不能完全恢复原始数据，但是所损失的部分对理解原始图像的影响缩小，却换来了大得多的压缩比，常见的有损压缩格式有JPEG、GIF、MP3、MP4等。无损压缩是利用数据的统计冗余进行压缩，可完全恢复原始数据而不引起任何失真，但压缩率是受到数据统计冗余度的理论限制，一般为2：1到5：1，这类方法广泛用于文本数据，程序和特殊应用场合的图像数据（如指纹图像，医学图像等）的压缩。题干中的压缩为有损压缩，故本题答案是：有损压缩。

答案有问题，这种压缩方法属于**无损压缩**。因为在压缩和解压缩过程中，图像的原始数据不会丢失或改变，能够完全恢复原始图像。

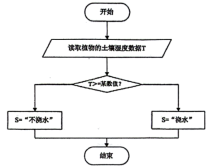
（3）软件的逻辑层次分为用户界面层、业务逻辑层、数据访问层，下列选项中，属于APP的业务逻辑层的是（ ）。

A．利用程序统计和分析老人的健康数据 B．用户登录界面

C．获取数据查询老人每日体征信息的程序 D．更新老人健康数据表的记录

（3）用户界面层是整个系统的最外层，和用户直接交互，用于显示数据和接收用户输入的数据。业务逻辑层是系统架构的核心，它处于数据访问层与表示层中间，起到了数据交换中承上启下的作用。数据访问层，其功能主要是负责数据库的访问，可以访问数据库系统、二进制文件、文本文档或是XML文档， 简单的说法就是实现对数据表的Select，Insert，Update，Delete的操作。故本题答案是：AC。

（3）小申在网上查询了植物的习性和传感器的参数，设计了一个算法来判断植物是否需要浇水，该算法流程图不满足算法特征的是（ ）。



A．有零个或多个输入 B．确定性 C．有一个或多个输出 D．有穷性

（3）算法的特征：有穷性（执行有限步后结束，且每一步的执行时间也是有限的）；确定性（每一步都有确切的含义）；输入（有零个或多个输入）；输出（至少产生一个输出）；可行性（在有限次运行后完成）。观察流程图，其中：“T>=某数值？”违背了算法的确定性；该算法没有输出，违背了算法的输出性，故本题答案是：BC。

## 2023育才

1. 关于信息的数字化，以下叙述正确的是（ ）

A. 各种形式的信息转换成不同进制存放在电脑中的过程称为信息的数字化

B. 磁盘中的影视信息在计算机上播放的过程就是信息的数字化

C. 计算机中存放的图像数据信息显示在显示器上的过程是信息的数字化

D. 将声音、影像等信息转化为二进制数形式的过程是信息的数字化

【详解】本题考查的是数字化。各种形式的信息转换成二进制存放在电脑中，选项A说法错误；磁盘中的影视信息已经是数字化了，选项B说法错误；计算机存放的图像数据，就已经是信息的数字化，选项C说法错误；将声音、影像等信息转化为二进制形式的过程是信息的数字化，选项D说法正确；故本题选D选项。

2. 2023年9月21日下午，“天宫课堂”第四课又开讲了，神舟十六号航天员景海鹏、朱杨柱、桂海潮在轨进行太空科普授课，三位博士带着全国青少年走进新教室——梦天实验舱，向同学们展示介绍梦天实验舱的工作生活场景。从信息与信息处理的角度分析中国空间站的“天宫课堂”授课活动，下列描述不正确的是（ ）

A. “天宫课堂”与地面之间的信息传输离不开载体

B. “天宫课堂”被广泛关注和学习，体现了信息的共享性

C. “天宫课堂”航天员与学生互动交流的信息没有经过数字化处理

D. “天宫课堂”课堂可以借助网络平台加快信息的传播速度

【详解】本题考查信息相关内容。信息的特征：（1）载体依附性：信息本身是抽象的，必须通过载体才能体现。载体就是承载信息的事物，语言、文字、声音、图像、视频等都可以作为信息的载体。同一信息可以依附不同的载体，例如同一个新闻事件，可以通过文字见诸于报纸，可以以声音的形式在广播播报，也可以视频的形式在电视中播放。（2）可加工性：信息可以加工和处理，例如统计用户的评价信息，确定商铺的信誉等级。（3）可存储性：信息可以脱离它所反映的事物被存储、保存和传播。（4）共享性：信息是可以传递和共享的，例如报纸上读到的新闻，再讲给同学听。（5）时效性：信息有一定的时效性，过了特定的时间信息就失去了效用，比如开奖后的彩票号码。（6）真伪性：信息有真也有假，我们需要提高辨别信息真伪的能力。（7）价值性：信息作为一种特殊的资源，具有相应的使用价值，能够满足人们某些方面的需要。ABD选项说法正确。C选项，“天宫课堂”航天员与学生互动交流的信息经过数字化处理，选项说法错误。故本题答案是C选项。

10. 公众号小编会定期将各种资料上传到云盘进行备份，为了节省云盘空间，他用Winrar软件对所有社团资料进行了压缩，以下关于数据压缩的叙述中，正确的是（ ）

A. 用Winrar压缩文件属于有损压缩

B. 文档数据不适合使用有损压缩

C. 图像数据从bmp格式转化为jpg格式属于无损压缩

D. 常见的音频数据有损压缩格式有wav和mp3

【详解】本题考查数据压缩技术相关内容。有损压缩格式是一种能够减小数据文件大小并且丢失一定质量格式，它通常被用于存储音频、视频、图像等多媒体文件，常见的有损压缩格式有JPEG、GIF、MP3、MP4等；无损压缩是指在压缩文件的同时不损失任何数据或质量，以达到减小文件大小的目的。A选项，用Winrar压缩文件属于无损压缩，选项错误。B选项，文档数据不适合使用有损压缩，选项正确。C选项，图像数据从bmp格式转化为jpg格式属于有损压缩，选项错误。D选项，WAV格式是实际声音的采样和编码，不是压缩格式，选项错误。故本题答案是B选项。

18. 数据如下图所示，现在要对其整理，以下说法正确的是（ ）page number 6



A. 第3行和第4行的数据完全一样，可以用dropna()来删除重复记录

B. 第6行中的温度38.2可能是异常值，一般使用箱形图或散点图能清晰地观察到异常值的存在

C. 第6行中的温度38.2可能是异常值，可以用前后两条记录中温度的平均值修正该异常值

D. 第7行有缺失值，在不研究湿度和风速的情况下，可以保留这条记录

【答案】B

【解析】

【详解】本题考查数据处理相关内容。A选项，在Python中,dropna()是一个Pandas库中的函数，用于从数据框（DataFrame）中删除包含缺失值(NaN)的行或列，它用于数据清洗和预处理阶段，以便去除缺失值，使数据更加规整，选项说法错误。B选项，散点图也叫X-Y图，它将所有的数据以点的形式展现在直角坐标系上，以显示变量之间的相互影响程度，点的位置由变量的数值决定，散点图能够展示数据的分布情况，通过观察数据点的密集程度和分布形态，推断出数据的离散程度和集中程度，散点图能帮助发现异常值或离群点，这些点在图上会与其他点有明显的偏离或不符合预期的位置，表示数据的异常情况或错误；箱形图是一种用作显示一组数据分散情况资料的统计图，因形状如箱子而得名，在各种领域也经常被使用，常见于品质管理，快速识别异常值。选项正确。C选项，出现异常值，应进行删除或修正，不能用前后两条记录中温度的平均值修正该异常值，选项错误。D选项，在处理缺失数据时，我们可以选择不同的重建方法，如删除缺失数据、均值填补、插值法、回归方法和机器学习方法，选项错误。故本题答案是B选项。

## 2022-2023奉贤

1. 中国城镇化数据分析研究

城镇化是人类社会发展的客观趋势，是国家现代化的重要标志，城镇化水平又称为城镇化率，是衡量城镇化程度的数量指标，也是衡量一个国家和一个地区社会经济发展水平的重要标志。中国正在加快推进新型城镇化建设，计划在2020年实施1亿左右农业转移人口和其它常住人口在城镇落户，常住人口城镇化率达到60%，户籍城镇化率达到45%。

（1）在研究某国城镇化率之前，需要先了解当前的人口数据。伴随着科学技术的发展，数据采集也变得越来越方便快捷，1953年我国进行第一次人口普查时，普查登记的实际时间历时半年多，而如今伴随着互联网等技术的应用，人口普查耗时越来越短，但是数据采集的基本流程是不变的，正确的数据采集流程为\_\_\_\_\_\_\_→\_\_\_\_\_\_\_→\_\_\_\_\_\_\_→\_\_\_\_\_\_\_（注：按顺序填写正确的步骤）。

①实施数据采集 ②明确数据要求 ③选择采集方法 ④确定数据来源

1. 正确的数据采集流程为：明确数据需求→确定数据来源→选择采集方法→实施数据采集。

2. 随着社会的高速发展，每天所产生的垃圾也越来越多，为了减轻环境压力，小申所住的小区里新安装了一台具有“用户识别”“自动称重”“满溢提醒”等功能的智能垃圾回收机，如图所示。居民投递可回收物时，能够根据自动称重的重量获得相应金额。这些回收物送到分拣工厂后经机器人精细分类，最终进入相应的再生产工厂变废为宝。



（1）智能垃圾回收系统的“满溢提醒”功能是借助超声波传感器来获取相关数据。这种数据获取的方式属于（ ）。（单选题）

A. 自动采集 B. 人工输入 C. 智能识别 D. 自动控制

（1）智能垃圾回收系统的“满溢提醒”功能是借助超声波传感器来获取相关数据。这种数据获取的方式属于自动采集，故选A选项

（6）第一空使用 for 循环遍历第 1 至第 x 天，循环变量是i。range(start, stop, [step])，start: 计数从 start 开始。默认是从 0 开始。例如range（5）等价于range（0， 5）；stop: 计数到 stop 结束，但不包括 stop。例如：range（0， 5） 是[0, 1, 2, 3, 4]没有5；step：步长，默认为1。例如：range（0， 5） 等价于 range(0, 5, 1)。因此填for i in range(1,x+1)

## 2022上海中学

（3）这款智能垃圾回收机有语音交互功能模块，方便用户进行功能选择、报修、意见收集等任务的实现。语音交互功能模块在工作时借助于内置的声2023.1线上音传感器，将采集到的\_\_\_\_\_\_\_\_\_\_\_信号转换为\_\_\_\_\_\_\_\_\_\_\_信号，即通过\_\_\_\_\_\_\_\_\_\_\_、\_\_\_\_\_\_\_\_\_\_\_、编码的过程，完成转换后才能让数字设备进行存储和处理（ ）。

A．数字 模拟 采样 量化 B．模拟 数字 量化 采样

C．模拟 数字 采样 量化 D．数字 模拟 量化 采样

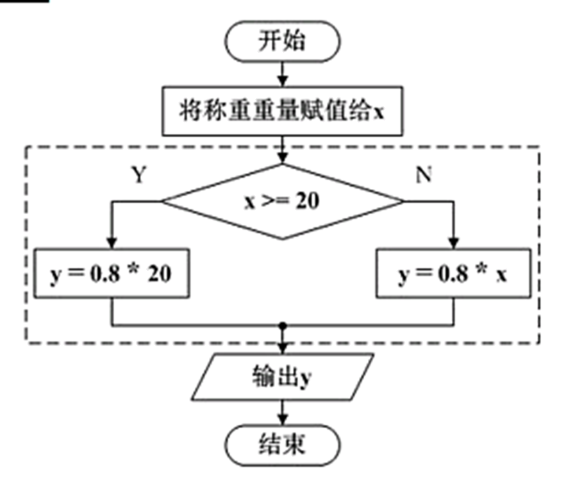
（3）将采集到的模拟信号转换为数字信号，即通过采样、量化、编码的过程，完成转换后才能让数字设备进行存储和处理，故选C选项。

（5）为了鼓励更多人使用智能垃圾回收机进行垃圾分类和回收，该机器可以对每次合理的分类投放给予一定的积分奖励。具体做法是：投放前可扫脸或输入手机号及密码进行登录。从数据安全防护的角度来看，输入手机号及密码的方式才能登录，是采用了\_\_\_\_\_\_\_\_\_\_\_方法进行了用户数据的保护；登录成功后个人的界面（如下图）则采用了\_\_\_\_\_\_\_\_\_\_\_方法进行了用户的数据安全防护。



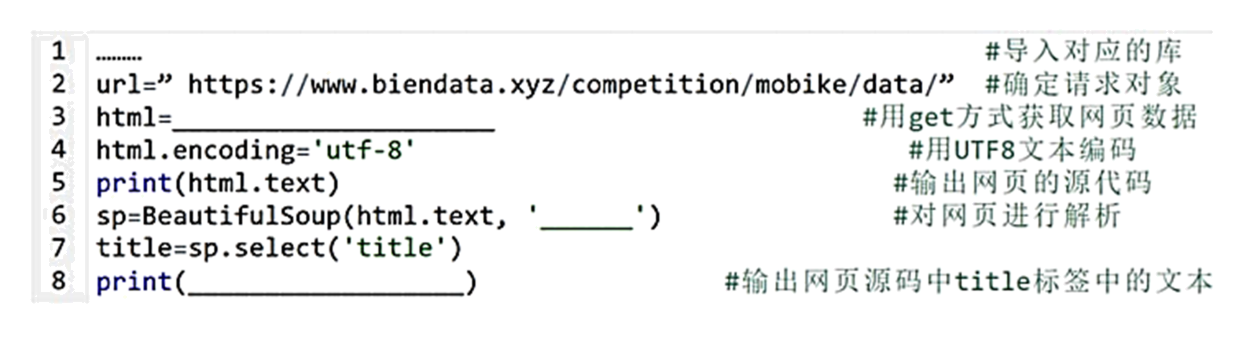
1. 从数据安全防护的角度来看，输入手机号及密码的方式才能登录，是采用了身份认证方法进行了用户数据的保护；登录成功后个人的界面（如下图）则采用了隐蔽技术方法进行了用户的数据安全防护，中间位用星号显示。

（10）小申用流程图设计了某次投递所获金额的算法（如下图），虚线框中算法的基本控制结构是\_\_\_\_\_\_\_\_\_\_\_，若可回收物的重量是20kg，表达式x>=20的值是\_\_\_\_\_\_\_\_\_\_\_。



1. 由图可知，若可回收物的重量是20kg，则满足判断条件，因此表达式x>=20的值是True。

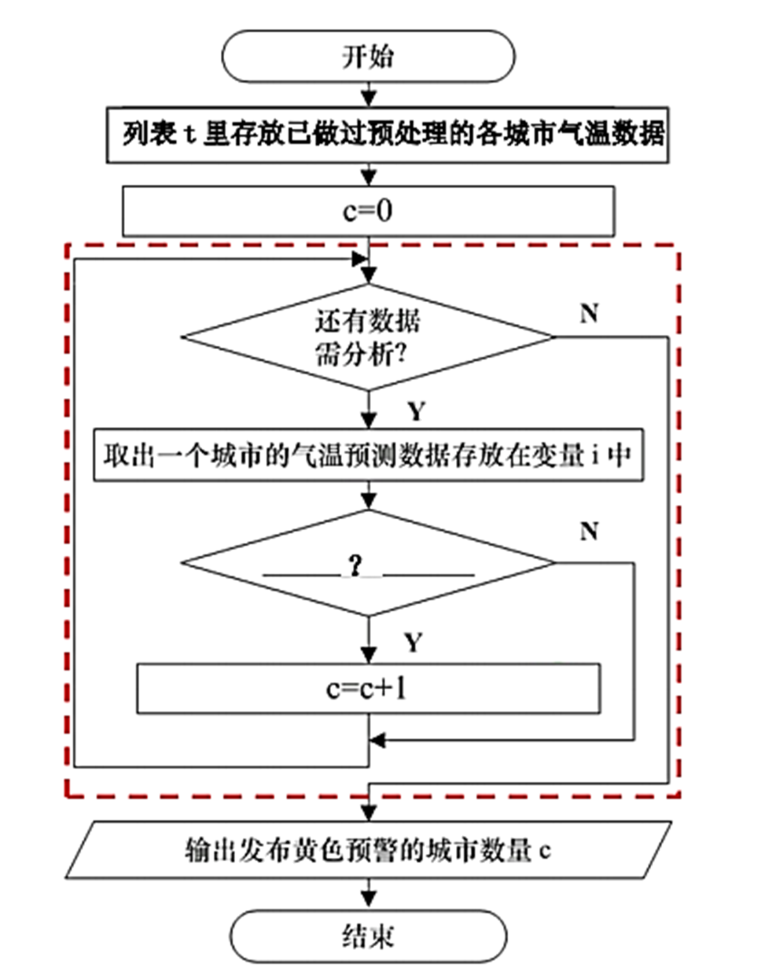
（3）网页请求方式有多种，GET是最常见的方式，一般用于获取或者查询资源信息，响应速度较快。若已经导入了方便发送HTTP请求的库，要用get方式获取网页数据的python代码（即下图第3行的横线上）需填写 requests . get ( url )；获取了网页信息后需要使用BeautifulSoup库中的方法进行网页解析，则代码中对应横线（图第6行）处应填写 html . parser



（3）要用 get 方式获取网页数据的 python 代码为 requests . get ( url )。其中 requests 为库名， get 为方法名。 url 是要抓取的网址。html . parser是Python内置的专门用来解析HTML的模块。利用html . parser我们可以分析出一段HTML里面的标签、数据等，是一种处理HTML的简便途径，因此第6行处填写 html . parser 。表示用 html 代码格式解析获得的信息。

（4）接上题，要输出网页源码中title标签的文本则先要用select属性抓取title标题，然后再使用print命令输出，那么print后的横线处（上图第8行）应填写\_\_\_\_\_\_\_\_\_\_\_

A．title．Text B．sp．text C．html．title．text D．sp．title．text

（4）title 表示获得的网页 title 部分。 title . Text 表示 title 标签中的内容。故本题选A选项。

③以上流程图中，虚线框内是典型的\_\_\_\_\_\_\_\_\_\_\_算法

以上流程图中，虚线框内是典型的多分支结构。

（7）为了统计出7月份低于平均日照时数的天数，小申打算编写Python程序解决。程序中sunlight用于记录“7月份每天的日照时数”的变量名，c是记录天数的变量名，i是循环控制的变量名，若要设置变量存放7月份日照总时数，从命名的规则来看，下面的变量名\_\_\_\_\_\_\_\_\_\_\_是不合法的

A．Import B．sum-days C．printJuly\_total D．Junetotal

（7）变量的命名规则有：只能以字母或下划线开头，不能包含特殊符号，不能是关键字。Import是关键字，sum-days含有特殊符号，因此均不能作为变量名，故本题选AB选项。