

深圳市第十三届职工技术创新运动会暨2023年深圳技能大赛— 人工智能训练师职业技能竞赛理论题库

序号	题干	题目类型	正确答案	答案选项A	答案选项B	答案选项C	答案选项D
001	人工智能的定义是什么？	单选	A	计算机模拟人类智能的技术	机器能够独立思考和学习	使机器具备人类的情感和意识	利用机器进行科学研究的方法
002	以下哪种技术不属于人工智能？	单选	D	机器学习	自然语言处理	数据挖掘	云计算
003	人工智能中的“机器学习”是指什么？	单选	B	让机器可以像人类一样学习知识	通过算法让机器从数据中学习和改进	让机器具备自主学习能力	通过神经网络让机器学习
004	人工智能中的“深度学习”是指什么？	单选	C	一种特殊的机器学习算法	让机器可以进行深度思考的能力	通过模拟人脑神经网络的方式进行学习	让机器可以进行复杂问题的推理和决策
005	以下哪个不是人工智能应用的例子？	单选	D	语音助手	机器人导游	自动驾驶汽车	手机游戏
006	人工智能中的“自然语言处理”是指什么？	单选	A	让机器可以理解和处理人类的自然语言	让机器能够像人类一样进行语音交流	利用机器进行语言学研究的方法	让机器能够进行自动翻译
007	人工智能中的“计算机视觉”是指什么？	单选	C	让机器可以理解和处理图像和视频	让机器能够像人类一样进行视觉感知	让机器可以进行图像识别和分析	利用机器进行图像处理和图像生成的技术
008	以下哪个不是人工智能的伦理问题？	单选	D	隐私保护	自主决策	就业岗位流失	网络安全
009	人工智能的发展会对社会产生哪些影响？	单选	D	提高生产效率和经济发展	减少人类工作岗位	改变人类生活方式和社会结构	所有选项都正确
010	人工智能的未来发展方向包括以下哪些？	单选	D	强化学习	量子计算	多模态学习	所有选项都正确
011	人工智能的历史可以追溯到哪个时期？	单选	A	20世纪50年代	19世纪80年代	21世纪90年代	18世纪30年代
012	以下哪个不是人工智能的子领域？	单选	B	机器学习	自动化	计算机视觉	自然语言处理
013	人工智能中的“强化学习”是指什么？	单选	B	让机器具备自主决策能力	通过奖惩机制让机器学习最优策略	利用神经网络进行学习和决策	让机器能够进行复杂问题的推理和决策
014	以下哪个不是人工智能中常用的算法？	单选	D	决策树	朴素贝叶斯	深度学习	遗传算法
015	人工智能中的“机器人”是指什么？	单选	A	一种可以执行任务的机器设备	一种具备智能和意识的机器	一种可以进行复杂推理的机器	一种可以进行学习和决策的机器
016	人工智能中的“专家系统”是指什么？	单选	A	一种可以模拟人类专家知识的系统	一种可以进行自主学习的系统	一种可以进行复杂推理的系统	一种可以进行图像识别的系统
017	人工智能中的“语音识别”是指什么？	单选	D	让机器能够理解和识别人类的语音	让机器能够进行语音合成	让机器能够进行语音翻译	让机器能够进行自动语音识别

018	机器学习是一种:	单选	A	人工智能技术	数据分析方法	编程语言	机械工程技术
019	在机器学习中, 监督学习是指:	单选	B	机器通过观察数据自主学习	机器根据给定的数据和标签进行学习	机器通过交互式学习改进自身性能	机器根据强化学习算法进行学习
020	以下哪项不是机器学习的主要任务?	单选	D	分类	聚类	回归	排序
021	机器学习的一个常见应用是:	单选	D	语音识别	图像处理	自然语言处理	所有上述
022	以下哪项不是机器学习中的评估指标?	单选	D	准确率	精确率	召回率	速度
023	什么是过拟合 (Overfitting) ?	单选	A	模型在训练集上表现良好, 但在测试集上表现差	模型在训练集上表现差, 但在测试集上表现良好	模型对新数据的泛化能力强	模型对新数据的泛化能力弱
024	以下哪种算法不属于监督学习算法?	单选	C	决策树	支持向量机	K均值聚类	神经网络
025	什么是梯度下降 (Gradient Descent) ?	单选	A	一种优化算法, 用于最小化损失函数	一种分类算法, 用于处理非线性数据	一种数据预处理技术, 用于降低数据维度	一种特征选择方法, 用于选择最相关的特征
026	智慧医疗的发展目前不受哪些因素影响	单选	C	缺少复合型人才	用户隐私数据安全问题	医学水平不够	市场主需求量不足以带动发展
027	人工智能核心技术体系不包括以下哪个层面()	单选	B	通用技术层	中间层	应用技术层	基础技术层
028	机器学习的流程包括:分析案例、数据获取、()和模型验证这四个过程	单选	C	数据清洗	数据分析	模型训练	模型搭建
029	机器翻译属于下列哪个领域的应用()	单选	A	自然语言系统	机器学习	专家系统	人类感官模拟
030	为了解决如何模拟人类的感性思维, 例如视觉理解、直觉思维、悟性等, 研究者找到一个重要的信息处理的机制是()	单选	B	专家系统	人工神经网络	模式识别	智能代理
031	下面关于函数说法错误的是()	单选	C	函数的三要素是函数名、参数和返回值, 定义函数时参数和返回值不是必须的	如果没有return语句, 则Python函数默认返回值为None	函数不可以嵌套	使用def定义函数
032	关于函数参数传递中, 形参与实参的描述错误的是()	单选	D	实参与形参分别存储在各自的内存空间中, 是两个不相关的独立变量	Python实行按值传递参数。值传递指调用函数时将常量或变量的值传递给函数的参数	实参与形参的名字可以相同, 也可以不同	在函数内部改变形参的值时, 实参的值也会改变的
033	已知列表对象ls, 哪个选项对ls.append(x)的描述是正确的()	单选	B	x不能是一个列表对象	只能向列表ls最后增加一个元素x	向列表ls最前面增加一个元素x	替换列表ls最后一个元素为x
034	专家系统中知识库的知识可以获取的来源是	单选	A	领域专家	专家系统的用户	计算机系统管理员	专家系统程序的开发者

035	关于Python的列表，以下选项中描述错误的是	单选	A	Python列表的长度不可变	Python列表是一个可以修改数据项的序列类型	Python列表用中括号[]表示	Python列表是包含0个或者多个对象引用的有序序列
036	下列选项不属于专家系统底层核心部分的是	单选	D	知识库	推理机	规则库	决策树
037	如果问题存在最优解，则下面几种搜索算法中，()必然可以得到该最优解	单选	A	广度优先搜索	深度优先搜索	有界深度优先搜索	启发式搜索
038	研究某超市销售记录数据后发现，买啤酒的人很大概率也会购买尿布，这种属于数据挖掘的那类问题	单选	A	关联规则发现	聚类	分类	自然语言处理
039	关于Python的字典类型，以下选项中描述错误的是()	单选	D	用{}标识,以key:value的形式存储数据	字典当中的元素是通过键(key)来存取的,而不是通过索引去取值	字典的元素是键值对,键(key)必须使用不可变类型	同一个字典中,键(key)不是唯一的
040	关于Python的字典类型，以下选项中描述正确的是()	单选	C	通过def函数创建字典	键值对之间以分号分隔	如果通过get(key)函数获取值:key不存在返回None	如果通过dict['key']获取值:key不存在不会报错
041	专家系统的发展趋势不包括	单选	D	知识库变大	推理引擎更加专用	用户接口更多样	用户需求量减少
042	BP神经网络模型拓扑结构不包括	单选	D	输入层	隐层	输出层	显层
043	机器学习研究如何通过计算的手段,利用经验来改善系统自身的性能,请问机器学习利用数据训练出什么	单选	A	模型	表结构	结果	报表
044	BP神经网络的学习规则是	单选	B	梯度上升法	梯度下降法	梯度提升法	梯度曲线法
045	关于Python的元组类型，以下选项中描述错误的是()	单选	A	元组中元素不可以是不同类型	元组一旦创建就不能被修改	Python中元组采用逗号和圆括号(可选)来表示	一个元组可以作为另一个元组的元素,可以采用多级索引获取信息
046	关于Python的元组类型，以下选项中描述正确的是()	单选	B	不可以索引取值	可以切片操作	元祖是可变序列	使用中括号包裹元素
047	机器学习的实质在于	单选	A	找	想	判断	理解
048	一个机器学习系统的基本结构不包括以下那个	单选	A	感知	学习	环境	知识库

049	所谓不确定性推理就是从()的初始证据出发,通过运用()的知识,最终推出具有一定程度的不确定性但却是合理或者近乎合理的结论的思维过程	单选	A	不确定性, 不确定性	确定性, 确定性	确定性, 不确定性	不确定性 确定性
050	机器通过人类发现的问题空间的数据,进行机器学习,具有在人类发现的问题空间中求解的能力,并且求解的过程与结果可以被人类智能() 此为机器智能的产生	单选	A	理解	参考	相同	采纳
051	第一个成功应用的专家系统是	单选	B	ELIZA	Dendral	Xcon	Deppblue
052	对抗学习中两个网络互相竞争,一个负责生成样本,另一个负责	单选	A	判别样本	计算样本	统计样本	生成样本
053	影响深度卷积神经网络算法的关键参数是	单选	B	网站数量	网络深度	网络结构	网络主体
054	深度学习的实质是	单选	B	推理机制	映射机制	识别机制	模拟机制
055	在深度优先策略中, open表是()的数据结构	单选	B	先进先出	先进后出	根据估价函数值重排	随机出
056	归纳推理是()推理	单选	B	从一般到个别	从个别到一般	从个别到个别	从一般到一般
057	JSON对象结构以"_____"开始,以"_____"结束	单选	A	{ }	()	[]	< >
058	JSON对象关键字和值之间以"_____"分隔	单选	C	/	,	:	.
059	JSON数组结构以"_____"开始,以"_____"结束	单选	C	{ }	()	()	< >
060	XML格式的文件由_____组成	单选	B	属性	标签对	键值对	逗号分隔的字符串
061	如下关于XML格式的文件,哪一个不是真的	单选	B	属性必须加引号	不区分大小写字母	有根元素	有关闭标签
062	Executable File是由_____执行的	单选	D	用户程序	应用程序	虚拟机	操作系统
063	CSV文件_____	单选	A	开头不留空不含列名	一行数据可跨行	用半角引号将字段值包含起来	一行数据里用冒号隔开
064	文本编码类型不可为	单选	C	ASCII	Unicode	FLV	UTF-8
065	在同一台计算机中,内存和外存存取速度	单选	C	看情况	一样快	内存快	外存快
066	32位微处理器中的32表示的技术指标是	单选	B	速度	字长	内存个数	接口数
067	Labelme不能对图像进行_____标注	单选	B	多边形	曲线	矩形	多段线
068	Labelme以_____文件存储标注信息	单选	D	XML	CSV	TXT	JSON

069	Labelme是一个图像标注软件，使用语言编写	单选	C	C	JAVA	Python	C++
070	视频源数据管理包含_____	单选	C	视频信息管理和视频档案管理	视频质量管理和视频目录管理	视频信息管理和视频目录管理	视频时间管理和视频字幕管理
071	ImageNet项目标注了_____图像	单选	C	万张图像	百万张图像	千万张图像	亿张
072	CIFAR-100中，每类各有_____个训练图像和_____个测试图像	单选	A	500, 100	100, 500	200, 300	300, 200
073	属于闭区域标注的有_____	单选	B	线标注和多边形标注	曲线标注和多边形标注	曲线标注和点标注	曲线标注和多边形标注
074	图像区域标注现在主要是	单选	A	手动标注	自动化标注	半自动化标注	人工智能标注
075	图像前处理对图像进行	单选	A	查重处理	尺寸处理	分辨率处理	标注
076	开区域标注常见的标注有	单选	D	曲线标注	多边形标注	框标注	线标注
077	交互式智能图像分割标注针对图像分割标注的预识别算法，实现通过机器对图像进行智能分割，再进行_____，从而大幅提高生产效率，减少人工成本，减少人工差错	单选	B	机器标注和筛选	人工修边和筛选	人工审核	自动识别与标注
078	老张交换名片后，小王打开手机中安装的灵云智能输入法 app，拍照老张的名片，很快得到名片文字信息并保存，这其中最主要应用的技术是	单选	A	模式识别	文字合成	图像搜索	图像还原
079	()是自然语言处理的重要应用，也可以说是最基础的应用。	单选	C	文本识别	机器翻译	文本分类	问答系统
080	下面()是LabelImg数据标注结果文件的后缀名。	单选	B	.avi	.xml	.jpg	.mov
081	flv格式是()的常见后缀名。	单选	B	语音文件	视频文件	文本文件	图像文件
082	在计算机内部，数据的表示形式是()。	单选	C	八进制	十进制	二进制	十六进制
083	将原始数据进行集成、变换、维度规约、数值规约是以下哪个步骤的任务()	单选	C	频繁模式挖掘	分类和预测	数据预处理	数据流挖掘
084	下面哪种不属于数据预处理方法()。	单选	D	变量代换	离散化	聚类	估计遗漏值
085	在Windows操作系统中，当双击myfile.docx文件时下面说法正确的是()	单选	C	直接在notepad下打开myfile.docx文件	直接在记事本下打开myfile.docx文件	直接在Word下打开myfile.docx文件	弹出选择打开方式对话框
086	CSV文件是以什么标点符号为分隔符的()。	单选	A	逗号	句号	单引号	双引号
087	下面哪种说法是错误的	单选	D	文本文件是基于字符编码方式存储文件	二进制文件是按二进制的编码方式来存放文件的	数据是以二进制编码方式存储在计算机文件中的	计算机只能识别文本文件

088	下面对数据标注描述哪一个是正确的	单选	D	所有的数据标注都可以使用脚本语言自动标注	可以通过算法来实现数据的自动标注	OCR手写转录完全可以通过识别工具实现自动转录	数据标注是个重复性很强的工作
089	下面对JSON文档叙述不正确的是()	单选	B	Object是一个无序的“‘名称/值’对”集合	一个对象以“[”开始，“]”结束	每个“名称”后跟一个“:”(冒号)	“‘名称/值’对”之间使用“,”分隔
090	下面对数据标注描述哪一个是正确的()	单选	D	所有的数据标注都可以使用脚本语言自动标注	可以通过算法来实现数据的自动标注	OCR手写转录完全可以通过识别工具实现自动转录	数据标注是个重复性很强的工作
091	下列关于合格数据标注员应该具备的素质中,描述错误的是	单选	C	持续的学习力是标注工作的基础	有耐心才能坚持在标注行业	具备一定的管理能力	较强专注力可提高个人工作效率
092	标注公司的岗位角色包括	单选	D	数据标注员	质检员	项目负责人	以上都是
093	下列关于数据标注员对数据进行标注的描述,错误的是	单选	C	需要参照数据标注规则	遇到数据标注规则模糊的问题需要及时询问	按照自己的想法猜测标注	数据标注员需要总结经验并及时汇报遇到的问题
094	初级数据标注员应该具备的职业机能包括() ①拥有一定的职业道德规范 ②行业基础知识掌握 ③具备基础业务能力 ④具备质检能力 ⑤具备团队管理能力	单选	A	①②③	③④⑤	②③④	①②⑤
095	中级数据标注员应该具备的职业机能包括() ①具备较高的职业道德和一定的从业经验 ②具备较高的业务能力和责任心 ③具备业务开发能力 ④具备简单质检能力 ⑤具备团队管理能力	单选	D	①②③⑤	①②④⑤	①②③④⑤	①②④⑤
096	数据标注规则的特点包括() ①数据标注规则需要一致 ②数据标注规则需要不断完善 ③需求方优先规则 ④质检优先规则 ⑤数据标注规则不能改变	单选	A	①②③④	②③④⑤	①③④⑤	①②④⑤
097	语音识别产品体系有四部分,下列哪项不是体系之一。()	单选	B	语音合成	语音播放	语音识别	语义理解
098	以下数据单位从小到大排列的顺序是()	单选	B	GB B KB	B KB MB	KB ZB PB	B MB KB

099	人工智能在某种意义上来说是人工+智能，以下那些需要人工做的工作（）	单选	D	数据采集	数据清洗	做标签	以上都对
100	图像去雾的应用场景不包括（）。	单选	D	视频监控	远程感应	自动驾驶	美颜相机
101	不属于图像分类技术应用领域的是（）	单选	D	商品自动分类	运输车辆识别	残次品自动分类	自动驾驶技术
102	语音识别、语义理解、语音合成、OCR 识别、人脸识别等都属于 人工智能技术，以下哪种事物没有应用到人工智能技术（）	单选	D	机器猫	无人机	微信客服	哈利波特的扫帚
103	（）是人以自然语言同计算机进行交互的综合性技术，结合了语言学、心理学、工程、计算机技术等领域的知识。	单选	B	情感交互	语音交互	体感交互	脑机交互
104	通常一个标注任务的准确率高于（）称该任务合格。	单选	D	60%	80%	90%	95%
105	下列选项中，不属于生物特征识别技术的是（）	单选	C	步态识别	声纹识别	文本识别	虹膜识别
106	计算机视觉与很多学科都有密切关系，例如（）为计算机视觉提供了图像处理的方法与工具。	单选	A	数字图像处理	美术学	计算机绘图	物理学
107	（）致力于计算机理解人类语言的句子或词语，目的是实现人机交互。	单选	D	文字校对技术	自然语言生成技术	中文自动分词技术	自然语言处理技术
108	对于神经退行性疾病，（）将成为主要的检测手段。	单选	D	窥镜	计算机断层扫描	X光	核磁共振成像
109	我国在语音语义识别领域的领军企业是（）	单选	C	华为	图谱科技	科大讯飞	阿里巴巴
110	Cortana是（）推出的个人语音助手。	单选	C	苹果	亚马逊	微软	阿里巴巴
111	现有的主流人工智能对自然语言处理是基于（）视角的。	单选	C	第一人称	第二人称	第三人称	对象
112	仅个体变元被量化的谓词称为（）。	单选	A	一阶谓词	原子公式	二阶谓词	全称谓词
113	不是自然语言处理要实现的目标的是（）	单选	C	理解别人讲的话	对自然语言表示的信息进行分析概括或编辑	欣赏音乐	机器翻译
114	客流分析可准确分析顾客年龄、性别等信息，区别新老客户，助力精准营销，其应用的主要技术不包括（）	单选	C	人脸识别	比对	文字识别	搜索技术
115	观察一批数据中的每一个数据在所有数据的总和中所占的比例，适合的图形是（）	单选	D	散点图	柱形图	直方图	饼图
116	自然语言理解(Natural Language Understanding)是一项（）	单选	D	网络技术	语言技术	理解能力	人工智能技术
117	文本比语音中的内容多了一个字，应该怎么办？	单选	B	判该句无效	删除文本中多余的字	提交反馈“语音有误”	无需理会

118	若规范要求，语句前后，需要各预留0.3~0.5秒的静音段。你在句尾之后的0.4秒处，发现有个按键音，以下哪种方法最合适？	单选	C	在句子后0.5秒处切音	在句子后0.4秒处切音	在句子后0.3秒处切音	判该句无效
119	文本内容是：我饿了。但是录音人有点儿结巴，连着说了四个我。该怎么标注？	单选	C	判该句无效	四个我饿了	我我我我饿了	我饿了
120	一句语音中，出现后面有第二人说话的声音。该怎么办？	单选	A	判该句无效	只写出主说话人的内容	两人的说话内容都写出	只写出第二人的内容
121	可以标为坏数据的是（）。	单选	C	图片显示方向不正确	图片局部显示不全	图片模糊放大后仍看不清楚	图片不够清晰但放大后可以看清楚
122	一道大题内包含选择题和填空题，则该大题的属性是（）。	单选	C	选择题	填空题	其他	噪声
123	下列选项中没有体现人工智能在消费品与电商领域行业的应用的是（）。	单选	B	智能店铺管理	智能客服	无人超市	天猫精灵
124	下列选项中有关人工智能可以带来的商业价值分析不正确的是（）。	单选	A	在金融行业，通过人工智能技术在风险评估、资产配置、智能交易等方向的应用，预计人工智能将带来约6000亿元人民币的降本增益效益。	在零售行业，人工智能在推荐系统上的运用将提高在线销售的销量表现，同时更加精准的市场预测将降低库存成本，预计人工智能技术将带来约4200亿元人民币的降本与增益价值。	在医疗行业，通过人工智能技术在药物研发领域提高成功率、在医疗服务机构内提供疾病诊断辅助、疾病监护辅助等提高服务效率的应用，预计人工智能可以带来约4000亿元人民币的降本价值。	在汽车行业，人工智能在自动驾驶上的技术突破将带来约5000亿元人民币的价值增益。
125	以下不属于无人超市采用的智能技术的是（）。	单选	D	技术机视觉	深度算法学习	传感器定位	图像处理
126	下列说法中对通用人工智能说明错误的是（）。	单选	B	具备知识技能迁移能力，可以快速学习。	是真正在这次人工智能浪潮中起到影响的主角。	是众多科幻作品中颠覆人类社会的人工智能形象。	充分利用已掌握的技能来解决新问题、达到甚至超过人类智慧的人工智能
127	下列说法中对专用人工智能理解正确的是（）。	单选	A	在某一个特定领域应用的人工智能。	不是真正在这次人工智能浪潮中起到影响的主角。	充分利用已掌握的技能来解决新问题、达到甚至超过人类智慧的人工智能。	是众多科幻作品中颠覆人类社会的人工智能形象。

128	下列有关人工智能说法错误的是()。	单选	A	不是人的智能, 虽然不能像人那样思考、按时有可能会超过人的智能。	人工智能可以对人的意识、思维的信息过程的模拟。	它是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。	人工智能就是研究如何使计算机去做过去只有人才能做的智能工作。
129	下列说法中哪个选项反应了人工智能学科的基本思想和基本内容()。	单选	B	人工智能是计算机科学的一个分支, 它企图了解智能的实质, 并生产出一种新的能以人类智能相似的方式做出反应的智能机器。	人工智能就是研究如何使计算机去做过去只有人才能做的智能工作。	人工智能不是人的智能, 但能像人那样思考、也可能超过人的智能。	以上都不正确
130	人工智能是一门利用计算机模拟人类智能行为科学的统称, 它涵盖了训练计算机使其能够完成()等人类行为的范畴。	单选	D	自主学习、判断、执行	决策、判断、执行	自主学习、决策、执行	自主学习、判断、决策
131	()是汽车人工智能领域目前最为火热的方向。	单选	B	整车的智能营销	驾驶辅助系统	零部件的预测维修	数据驱动的产品优化
132	下列哪个选项是专用人工智能的定义()。	单选	A	在某一个特定领域应用的人工智能。	指具备知识技能迁移能力, 可以快速学习, 充分利用已掌握的技能来解决新问题、达到甚至超过人类智慧的人工智能。	是一门利用计算机模拟人类智能行为科学的统称。	是计算机科学的一个分支。
133	下列选项中哪些是不属于专用人工智能和通用人工智能之间的区别()。	单选	C	知识技能的迁移能力	跨领域推理能力	实现特定领域的应用	意识的认识与掌握
134	()是零售行业内应用最为广泛、效果最为显著的人工智能技术, 线上线下的零售巨头都在运用此技术帮助进行交叉销售、向上销售、提高复购率。	单选	D	智能店铺管理	无人超市	智能交易策略	精准营销与个性化推荐系统
135	淘宝美工——鲁班这一案例属于人工智能产业应用的哪个场景()。	单选	B	人工智能在汽车行业中的应用	人工智能在消费品和电商领域行业的应用	人工智能在金融行业的应用	人工智能在个人生活娱乐学习方面的应用
136	智能医疗可以在()等方面发挥重要作用。	单选	A	辅助诊疗、疾病预测、医疗影像辅助诊断、药物开发	货物搬运、仓储管理、货物配送	人脸识别、指纹解锁、产品检测	远程维护、企业管理、个性化定制

137	下面活动中应用了智能语音技术的是()。	单选	C	当天黑了家里的窗帘自动拉合	下雨天窗户自动关闭	夜间起床说声“开灯”，夜灯就亮了	当宝宝大声哭泣，智能床铃就开始播放音乐
138	通过交通信息采集系统采集道路中的车辆流量、行车速度等信息，经智能系统分析后调整各路口红绿灯时长属于人工智能在()领域的应用。	单选	D	智能物流	智能安防	智能控制	智能交通
139	下面不是应用于智能物流领域的机器人是()。	单选	A	迎宾机器人	搬运机器人	货架穿梭车	分拣机器人
140	下面关于人工智能说法错误的是()。	单选	B	人工智能一定程度上能模仿人类的活动	人工智能可以全面取代人类活动	人工智能的发展将改变人类社会生活	人工智能的发展也会产生一些负面的社会影响
141	下面关于企业客服机器人说法正确的是()。	单选	C	企业客服机器人是闲聊机器人	企业客服机器人能回答所有用户提出的问题	企业客服机器人属于聊天机器人	企业客服机器人不是以任务目的为导向的机器人
142	下面不属于领域知识智能问答机器人的是()。	单选	C	淘宝智能客服机器人	南航智能客服	分拣机器人	电信智能客服
143	在与智能客服机器人交流时，人们采用的是()。	单选	B	关键字	自然语言	纯语音	纯文字
144	不属于当今的人工智能发展方向的是()	单选	D	以机器与人结合而成的增强型混合智能系统	用机器、人、网络结合成新的群智系统	用机器、人、网络和物结合成的更加复杂的智能系统	用计算机模拟人工智能
145	生物特征识别技术是指通过个体生理特征或行为特征对个体身份进行识别认证的技术。下列技术不属于生物特征识别技术的是()。	单选	B	人脸识别	3D识别	虹膜识别	声纹识别
146	根据机器智能水平由低到高，()是正确的。	单选	A	计算智能、感知智能、认知智能	计算智能、感应智能、认知智能	机器智能、感知智能、认知智能	机器智能、感应智能、认知智能
147	三大流派的演化正确的是()。	单选	C	符号主义->知识表示->机器人	联结主义->控制论->深度学习	行为主义->控制论->机器人	符号主义->神经网络->知识图谱
148	()不是手机里常用的智能APP。	单选	C	美颜	语音助手	人脸识别	机器翻译
149	掀起人工智能发展的第三个高潮是由()的。	单选	B	计算驱动	数据驱动	知识驱动	常识驱动
150	只专注于完成某个特别设定的任务的人工智能属于()。	单选	B	超人工智能	弱人工智能	强人工智能	认知智能

151	()系统包括了学习、语言、认知、推理、创造和计划,目标是使人工智能在非监督学习情况下处理前所未见的细节,并同时与人类开展交互式学习。	单选	C	超人工智能	弱人工智能	强人工智能	认知智能
152	Python中的变量var如果包含字符串的内容,下面哪种数据类型不可能创建var?()	单选	C	list	string	char	dict
153	Python程序在执行一次之后会自动生成扩展名为()的字节码文件,以提高运行效率。	单选	D	.pyf	.pyb	.py	.pyc
154	Python函数定义的关键字是()	单选	A	def	function	import	main
155	下面python程序的运行结果是() a=10; def setNumber(): a=100; setNumber(); print(a)	单选	C	100	10100	10	10010
156	关于python列表,下列选项中,描述错误的选项是。	单选	D	列表中的元素之间用逗号分隔	列表中的元素可以是列表类型	列表中可以包含数字类型的元素	列表中的元素类型必须相同
157	python中,已知ls=[12,34.5,True,'test',3+5j],则下列选项中,输出结果为“['test']”的选项是()	单选	C	ls[3]	ls[4]	ls[3:4]	ls[4:5]
158	下列哪种说法是错误的()	单选	B	python是一门面向对象的语言;	python是一门面向过程的语言;	python是一种解释型语言的计算机程序设计语言;	python程序无需编译成二进制代码,而是在执行时对语句一条一条编译。
159	关于函数参数传递中,形参与实参的描述错误的是()。	单选	C	实参与形参分别存储在各自的内存空间中,是两个不相关的独立变量	Python实行按值传递参数。值传递指调用函数时将常量或变量的值传递给函数的参数	实参与形参的名字必须相同	在函数内部改变形参的值时,实参的值一般是不会改变的
160	Python脚本的扩展名是()	单选	B	.python	.py	.pt	.pg
161	关于 a or b描述错误的是()。	单选	C	若a= True b= True 则a or b==True	若a= True b= False 则a or b==True	若a= True b= True 则a or b==False	若a= False b= False 则a or b==False
162	关于函数说法错误的是()。	单选	B	函数可以没有参数	函数可以有多个返回值	函数可以没有return语句	函数都有返回值
163	以下不合法的pythhon变量名是()。	单选	B	Python2	N.X	sum	Hello World

164	关于python类 说法错误的是()。	单选	B	类的实例方法必须创建对象后才可以调用	类的实例方法必须创建对象前才可以调用	类的类方法可以用对象和类名来调用	类的静态属性可以用类名和对象来调用
165	list(range(6))[::2]的执行结果为()?	单选	C	[0, 1, 2, 3, 4, 5]	[2, 3, 4, 5, 6]	[0, 2, 4]	[1, 3, 5]
166	表达式[x for x in range(10) if x%2==0]的结果是()	单选	D	语法错误	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]	(2, 4, 6, 8, 10)	[0, 2, 4, 6, 8]
167	下面哪个数是float类型()?	单选	D	10	TRUE	1+2j	3.14
168	一个字符串str="Hello Python",则str[1:8]是()?	单选	B	Hello Py	ello Py	ello Pyt	Hello P
169	下面关于Python语言特点的描述, 错误的是()?	单选	C	Python语言是开源、跨平台的语言。	Python语言具有可移植性。	Python语言是FLOSS之一, 可以使用但不能随意进行改动。	Python语言是解释性语言, 通过解释器将源码转换成字节码。
170	输出函数print()参数中使用下面哪个属性来指定末尾符号为换行符()?	单选	B	sep=''	end='\n'	file=sys.stdout	flush=False
171	下面转义字符中哪个符号是制表符()?	单选	A	\t	\n	\\	\'
172	range(5)的取值范围是()?	单选	C	[0, 1, 2, 3, 4, 5]	[1, 2, 3, 4, 5]	[0, 1, 2, 3, 4]	[5]
173	下面哪个函数用于从键盘接收信息()?	单选	C	int()	print()	input()	range()
174	Python程序使用缩进来组织代码的层次结构, 1个缩进是几个空格()?	单选	C	1	2	8	8
175	关于Python变量, 说法正确的是	单选	B	变量无需赋值便可以直接使用	变量无需声明数据类型便可以直接赋值	Python变量只有数字型和字符串型2种	Python变量与其他所有高级程序设计语言变量的数据类型没有区别
176	在Anaconda中进行第三方库的安装, 正确的命令是	单选	A	pip install 包名	conda 包名	conda setup 包名	pip setup 包名
177	下面关于Python语言的说法, 错误的是	单选	C	Python源代码区分大小写	Python语言是解释性的, 可以在>>>提示符下交互输入Python语句	python语言是编译执行的, 不支持逐条语句执行方式	Python用#引出行注释
178	已有变量x和y, 以下()不能实现交换变量x和变量y的值。	单选	A	x=y; y=x	x, y=y, x	t=y; y=x; x=t	x=y+x; y=x-y; x=x-y
179	以下选项中, Python代码的注释使用的符号是()	单选	D	//	/*..... */	%	#
180	可以使用()接受用户的键盘输入。	单选	B	input命令	input()函数	int()函数	format()函数

181	已知area=1963.4375000000002, 执行print("{:.2f}".format(area))语句, 输出结果为	单选	D	1919	1963	1963.43	1963.44
182	下列数据类型中, ()属于无序数据类型。	单选	D	set、tuple	str、list	list、tuple	set、dict
183	设有变量a="Me", "You", 则变量a属于	单选	B	字符串	元组	列表	集合
184	数据表是由行(记录)和列(字段)构成, 因此也称()	单选	B	数据	二维表	表格	大数据
185	在数据表中, 表的“行”称为()	单选	C	数据	字段	记录	大数据
186	()是指对客观事件记录的符号, 是对客观事物的性质、状态及相互关系等进行记载的物理符号或这些物理符号的组合。	单选	A	数据	字段	记录	大数据
187	在数据表中, 表的“列”称为()	单选	B	数据	字段	记录	大数据
188	数据采集器不能采集以下哪种数据()	单选	D	图片	文字	网站信息	专利
189	下列哪种不是数据采集器	单选	A	Python爬虫	八爪鱼采集器	火车头采集器	后羿采集器
190	某用户在使用支付宝绑定银行卡时, 网站要求验证用户的真实姓名和身份证号码。这时要求采集的数据必须有()	单选	A	准确性	完整性	一致性	相关性
191	数据的存储方式有()	单选	D	Excel	CSV	数据库	以上都对
192	通用网络爬虫基本工作流程包含()	单选	D	抓取网页	数据存储	预处理	以上都对
193	防采集措施不包含以下哪种()	单选	C	封IP	验证码识别	不允许打开网页	登录
194	商务数据的来源不包括()	单选	A	个人数据	电子商务平台	社交平台	020数据
195	数据的获取途径没有()	单选	A	从别人数据库窃取	产品自有数据	调查问卷	互联网数据导入
196	常见的爬虫语言不包含()	单选	A	HTML	Python	java	php
197	以下对字典的说法不正确的是()	单选	D	字典可以为空	字典的键不能相同	字典的键不可变	字典的值不可变
198	如果把知识按照作用来分类, 下述()不在分类的范围内。	单选	B	用控制策略表示的知识, 即控制性知识。	可以通过文字、语言、图形、声音等形式编码记录和传播的知识, 即显性知识	用提供有关状态变化、问题求解过程的操作、演算和行动的知识, 即过程性知识。	用提供概念和事实使人们知道是什么的知识, 即陈述性知识。
199	下述()不是知识的特征。	单选	A	复杂性和明确性	进化和相对性	客观性和依附性	可重用性和共享性

200	人类智能的特性表现在哪4个方面。()	单选	B	聪明、灵活、学习、运用。	能感知客观世界的信息、能对通过思维对获得的知识进行加工处理、能通过学习积累知识增长才干和适应环境变化、能对外界的刺激作出反应传递信息。	感觉、适应、学习、创新。	能捕捉外界环境信息、能够利用利用外界有利因素、能够传递外界信息、能够综合外界信息进行创新思维。
201	人工智能的目的是让机器能够()，以实现某些脑力劳动的机械化。	单选	D	具有智能	和人一样工作	完全代替人的大脑	模拟、延伸和扩展人的智能
202	下列关于人工智能的叙述不正确的是()。	单选	C	人工智能技术它与其他科学技术相结合极大地提高了应用技术的智能化水平。	人工智能是科学技术发展的趋势。	因为人工智能的系统研究是从上世纪五十年代才开始的，非常新，所以十分重要。	人工智能有力地促进了社会的发展。
203	人工智能研究的一项基本内容是机器感知。以下列()不属于机器感知的领域。	单选	C	使机器具有视觉、听觉、触觉、味觉、嗅觉等感知能力。	让机器具有理解文字的能力。	使机器具有能够获取新知识、学习新技巧的能力。	使机器具有听懂人类语言的能力
204	尽管人工智能学术界出现“百家争鸣”的局面，但是，当前国际人工智能的主流派仍属于：()	单选	B	连接主义	符号主义	行为主义	经验主义
205	被誉为国际“人工智能之父”的是：()	单选	A	图灵(Turing)	费根鲍姆(Feigenbaum)	傅京孙(K. S. Fu)	尼尔逊(Nilsson)
206	人工智能的含义最早由()于1950年提出，并且同时提出一个机器智能的测试模型。	单选	C	明斯基	扎德	图灵	冯·诺依曼
207	下列哪个不是人工智能的研究领域()。	单选	D	机器证明	模式识别	人工生命	编译原理
208	AI是()的英文缩写。	单选	B	Automatic Intelligence	Artificial Intelligence	Automatic Information	Artificial Information
209	为了解决如何模拟人类的感性思维，例如视觉理解、直觉思维、悟性等，研究者找到一个重要的信息处理的机制是()。	单选	B	专家系统	人工神经网络	模式识别	智能代理

210	下述（ ）不是人工智能中常用的知识格式化表示方法。	单选	D	框架表示法	状态空间表示法	语义网络表示法	形象描写表示法
211	关于“与/或”图表示法的叙述中，正确的是（ ）。	单选	D	“与/或”图就是用“AND”和“OR”连续各个部分的图形，用来描述各部分的因果关系。	“与/或”图就是用“AND”和“OR”连续各个部分的图形，用来描述各部分之间的不确定关系。	“与/或”图就是用“与”节点和“或”节点组合起来的树形图，用来描述某类问题的层次关系。	“与/或”图就是用“与”节点和“或”节点组合起来的树形图，用来描述某类问题的求解过程。
212	神经网络研究属于下列（ ）学派。	单选	B	符号主义	连接主义	行为主义	都不是
213	已知初始问题的描述，通过一系列变换把此问题最终变为一个子问题集合；这些子问题的解可以直接得到，从而解决了初始问题。这是知识表示法叫（ ）。	单选	B	状态空间法	问题归约法	谓词逻辑法	语义网络法
214	在公式中 $\forall xP(x, y)$ ，存在量词是在全称量词的辖域内，我们允许所存在的 x 可能依赖于 y 值。令这种依赖关系明显地由函数所定义，它把每个 y 值映射到存在的那个 x 。这种函数叫做（ ）。	单选	B	依赖函数	Skolem函数	决定函数	多元函数
215	$A(AB)A$ 称为（ ）。	单选	C	结合律	分配律	吸收律	摩根律
216	$\sim(AB)\sim A\sim B$ 称为（ ）。	单选	D	结合律	分配律	吸收律	摩根律
217	如果问题存在最优解，则下面几种搜索算法中，（ ）必然可以得到该最优解。	单选	A	广度优先搜索	深度优先搜索	有界深度优先搜索	启发式搜索
218	训练图像分类模型时，对于图像的预处理，下列技术哪项经常要用（ ）。	单选	A	图像增强	图像灰度化	图片二值化	图片RGB 通道转换
219	语音识别属于人工智能中的（ ）。	单选	C	指纹识别研究范畴	数字识别研究范畴	模式识别研究范畴	字符识别研究范畴
220	下列传感器中，哪个不属于本体感知型传感器（ ）。	单选	D	陀螺仪	光电编码器	霍尔效应编码器	超声波测距传感器
221	下列哪个不是人工智能的研究领域（ ）。	单选	D	机器证明	模式识别	人工生命	编译原理
222	语音识别的定义（ ）。	单选	C	语音识别是获取语音的过程及技术手段	语音识别是识别说话人的技术手段	语音识别技术是识别语音中声学特征信息，提取及应用的过程。	语音识别是模仿人类听觉的技术手段
223	BP神经网络的学习规则是（ ）	单选	B	梯度上升法	梯度下降法	梯度提升法	梯度曲线法
224	视觉slam 中，下面哪一项不属于视觉传感器的分类（ ）。	单选	D	单目相机	多目相机	RGBD 相机	单反相机
225	根据科学流行定义，人工智能就是和人类（ ）相似的计算机程序。	单选	C	思考方式	表达方式	行为方式	外观外貌

226	下列哪一项不是常见的机器学习模型正则化方法	单选	A	数据优化	数据增强	引入参数范数惩罚项	模型集成
227	FOIL是()的学习算法	单选	A	一阶规则	序贯覆盖	命题规则	剪枝优化
228	Inception模块可以并行执行多个具有不同尺度的卷积运算或池化操作, 下列网络运用Inception的是()	单选	C	VGG	fast-RCNN	GoogLeNet	faster-RCNN
229	在不考虑标记样本时, 支持向量机试图找到()间隔划分超平面	单选	D	最小	最短	最长	最大
230	以下哪种卷积神经网络的设计引入了残差网络结构	单选	C	LeNet	GoogLeNet	ResNets	AlexNet
231	自然语言中的词语需要转化为计算机可以记录处理的数据结构, 通常会把自然语言中的词语转化为以下哪种数据结构	单选	B	标量	向量	结构体	有向图
232	下列哪项属于集成学习()	单选	C	决策树模型	K-means	Adaboost	kNN分类
233	我国劳动法的调整对象是()	单选	D	劳动关系	劳动关系和社会保障关系	劳动关系和社会保险关系	劳动关系以及与劳动关系有密切联系的其他关系
234	我国最早的劳动立法是北洋政府1923年公布的()	单选	B	《劳动立法原则》	《暂行工厂规则》	《劳动法案大纲》	《劳动法典》
235	《中华人民共和国劳动法》规定, 延长工作时间每月最长的时限是	单选	C	12小时	24小时	36小时	48小时
236	因工致残依劳动能力丧失的程度不同, 分为()	单选	D	4级	6级	8级	10级
237	根据《劳动法》规定, 国家确定职业分类, 对规定的职业制定职业技能标准, 实行()	单选	A	职业资格证书制度	学历文凭制度	培训证书制度	技能分类制度
238	养老保险关系属于()	单选	B	劳动关系	与劳动关系有密切联系的社会关系	国家进行劳动力管理中的关系	处理劳动争议过程中发生的关系
239	企业实行民主管理的基本形式是()	单选	B	工会	职工代表大会	企业管理委员会	股东大会
240	领取失业救济金的期限最长为()	单选	C	6个月	12个月	24个月	36个月
241	当事人申请劳动争议仲裁的时效是, 从知道或应当知道权利受侵害之日起()	单选	C	15日内	30日内	60日内	90日内
242	什么是K折交叉验证(K-fold Cross Validation)?	单选	C	一种无监督学习方法	一种特征选择技术	一种模型评估方法	一种数据预处理技术

243	以下哪项不是决策树算法的优点?	单选	D	易于理解和解释	可处理非线性数据	不需要数据预处理	可以处理大规模数据集
244	什么是特征工程 (Feature Engineering) ?	单选	C	一种无监督学习方法	一种数据可视化技术	一种数据预处理技术	一种优化算法
245	以下哪项不是神经网络的常见层类型?	单选	C	卷积层	池化层	随机森林层	全连接层
246	什么是深度学习 (Deep Learning) ?	单选	D	一种无监督学习方法	一种集成学习技术	一种机器学习算法	一种神经网络的扩展
247	以下哪项不是监督学习中的回归算法?	单选	C	决策树回归	支持向量机回归	K均值聚类	线性回归
248	什么是主成分分析 (Principal Component Analysis) ?	单选	D	一种无监督学习方法	一种特征选择技术	一种数据预处理技术	一种降维算法
249	以下哪项不是聚类算法的常见类型?	单选	C	K均值聚类	DBSCAN	随机森林	层次聚类
250	什么是支持向量机 (Support Vector Machine) ?	单选	C	一种无监督学习方法	一种特征选择技术	一种分类算法	一种回归算法
251	以下哪项不是聚类算法的评估指标?	单选	A	准确率	精确率	召回率	轮廓系数
252	什么是朴素贝叶斯 (Naive Bayes) ?	单选	C	一种无监督学习方法	一种特征选择技术	一种分类算法	一种回归算法
253	以下哪项不是强化学习的一个关键概念?	单选	D	奖励	状态	动作	标签
254	什么是随机森林 (Random Forest) ?	单选	C	一种无监督学习方法	一种特征选择技术	一种分类算法	一种回归算法
255	以下哪项不是强化学习算法?	单选	B	Q学习	卷积神经网络	蒙特卡洛方法	DQN
256	什么是逻辑回归 (Logistic Regression) ?	单选	C	一种无监督学习方法	一种特征选择技术	一种分类算法	一种回归算法
257	以下哪项不是逻辑回归的优缺点?	单选	B	易于解释	可以处理非线性数据	计算效率高	对异常值敏感
258	什么是卷积神经网络 (Convolutional Neural Network) ?	单选	C	一种无监督学习方法	一种特征选择技术	一种分类算法	一种回归算法
259	以下哪项不是卷积神经网络的组成部分?	单选	D	卷积层	池化层	全连接层	朴素贝叶斯层
260	什么是递归神经网络 (Recurrent Neural Network) ?	单选	C	一种无监督学习方法	一种特征选择技术	一种分类算法	一种回归算法
261	以下哪项不是递归神经网络的特点?	单选	D	可以处理序列数据	可以处理变长输入	具有记忆能力	只能处理静态数据
262	什么是生成对抗网络 (Generative Adversarial Network) ?	单选	D	一种无监督学习方法	一种特征选择技术	一种分类算法	一种生成模型
263	以下哪项不是生成对抗网络的组成部分?	单选	C	生成器	判别器	编码器	解码器
264	深度学习是一种机器学习的子领域, 其特点是:	单选	A	使用深层神经网络	使用深层贝叶斯模型	使用深度聚类算法	使用深度遗传算法
265	在深度学习中, 反向传播算法用于:	单选	A	训练神经网络	测试神经网络	初始化神经网络权重	正则化神经网络
266	深度学习使用的最常见的激活函数是:	单选	B	Sigmoid函数	ReLU函数	Tanh函数	Softmax函数
267	在深度学习中, 损失函数用于:	单选	A	衡量模型的性能	正则化模型	初始化权重	调整学习率

268	深度学习中常用的优化算法是:	单选	A	SGD (随机梯度下降)	K-means算法	KNN算法	决策树算法
269	在深度学习中, 批量归一化 (Batch Normalization) 的作用是:	单选	A	加速模型训练	减少模型复杂度	改善模型的泛化能力	防止模型过拟合
270	深度学习中的卷积神经网络 (CNN) 主要用于处理:	单选	A	图像数据	文本数据	时序数据	图像和文本数据
271	深度学习中的递归神经网络 (RNN) 主要用于处理:	单选	C	图像数据	文本数据	时序数据	图像和文本数据
272	深度学习中的生成对抗网络 (GAN) 主要用于:	单选	B	图像分类	图像生成	文本分类	文本生成
273	深度学习中的自编码器 (Autoencoder) 主要用于:	单选	C	特征提取	特征选择	特征降维	特征融合
274	深度学习中的注意力机制 (Attention) 主要用于:	单选	D	图像分类	图像生成	文本分类	文本生成
275	深度学习中的迁移学习 (Transfer Learning) 的目的是:	单选	B	提高模型的训练速度	提高模型的泛化能力	减少模型的参数数量	减少模型的复杂度
276	深度学习中的扩展方法GAN的全称是:	单选	B	Generative And Neural Networks	Generative Adversarial Networks	Gradient And Neural Networks	Gradient Adversarial Networks
277	深度学习中的LSTM是什么的缩写?	单选	A	Long Short-Term Memory	Long Sequential Training Model	Logistic Sequential Training Model	Logistic Short-Term Memory
278	深度学习中的Word2Vec是用来做什么的?	单选	A	词嵌入	图像分类	文本生成	特征选择
279	在深度学习中, Dropout的作用是:	单选	C	减少模型的参数数量	减少模型的复杂度	防止模型过拟合	加速模型训练
280	在深度学习中, 学习率衰减的作用是:	单选	D	加速模型训练	减少模型的参数数量	减少模型的复杂度	改善模型的泛化能力
281	深度学习中的图像分类任务中, 常用的数据增强方法有:	单选	D	随机裁剪	随机旋转	随机翻转	所有选项都对
282	在深度学习中, 滑动窗口 (Sliding Window) 是用来做什么的?	单选	B	图像分类	目标检测	文本分类	文本生成
283	在深度学习中, 数据集的划分常用的比例是:	单选	A	60%训练集, 20%验证集, 20%测试集	70%训练集, 15%验证集, 15%测试集	80%训练集, 10%验证集, 10%测试集	90%训练集, 5%验证集, 5%测试集
284	深度学习的目标函数 (Objective Function) 通常使用的是:	单选	A	交叉熵损失函数	均方误差损失函数	对数损失函数	KL散度损失函数
285	在深度学习中, 模型的容量 (Capacity) 是指:	单选	C	模型可以容纳的样本数量	模型的参数数量	模型的复杂度	模型的训练速度
286	深度学习中的模型正则化是为了:	单选	D	减少模型的参数数量	减少模型的复杂度	改善模型的泛化能力	防止模型过拟合

287	深度学习中的模型集成是指:	单选	A	将多个模型的预测结果进行融合	将多个模型的参数进行融合	将多个模型的训练样本进行融合	将多个模型的特征进行融合
288	深度学习中的模型压缩是指:	单选	A	减少模型的参数数量	减少模型的复杂度	改善模型的泛化能力	加速模型训练
289	深度学习中的模型蒸馏是指:	单选	A	将一个复杂模型的知识传递给一个简单模型	将一个简单模型的知识传递给一个复杂模型	将多个模型的知识进行融合	将多个模型的预测结果进行融合
290	深度学习中的模型评估常用的指标有:	单选	D	准确率	精确率	召回率	所有选项都对
291	深度学习中的模型调优常用的方法有:	单选	D	网格搜索	随机搜索	贝叶斯优化	所有选项都对
292	深度学习中的模型部署是指:	单选	A	将模型部署到生产环境中	将模型部署到测试环境中	将模型部署到开发环境中	将模型部署到训练环境中
293	深度学习中的模型解释性是指:	单选	A	可以解释模型的决策过程	可以解释模型的参数含义	可以解释模型的训练过程	可以解释模型的泛化能力
294	数据智能是指利用什么技术和方法来提取、分析和利用数据?	单选	D	人工智能	机器学习	数据挖掘	所有以上答案
295	数据仓库是指什么?	单选	A	存储数据的地方	数据分析的工具	数据可视化的技术	数据收集的过程
296	数据清洗是指什么?	单选	D	删除无用的数据	将数据转换成特定的格式	检查和修复数据错误	所有以上答案
297	数据挖掘是指什么?	单选	A	从数据中提取有用信息的过程	将数据转换成特定的格式	存储数据的地方	数据分析的工具
298	机器学习是指什么?	单选	A	计算机通过学习数据来改善性能的技术	将数据转换成特定的格式	存储数据的地方	数据分析的工具
299	数据可视化是指什么?	单选	A	通过图表和图形展示数据的过程	将数据转换成特定的格式	存储数据的地方	数据分析的工具
300	什么是监督学习?	单选	A	机器学习算法通过给定的输入和输出数据进行训练	机器学习算法通过观察和理解数据进行训练	机器学习算法通过试错方法进行训练	以上都不是
301	什么是无监督学习?	单选	B	机器学习算法通过给定的输入和输出数据进行训练	机器学习算法通过观察和理解数据进行训练	机器学习算法通过试错方法进行训练	以上都不是
302	什么是强化学习?	单选	C	机器学习算法通过给定的输入和输出数据进行训练	机器学习算法通过观察和理解数据进行训练	机器学习算法通过试错方法进行训练	以上都不是
303	什么是深度学习?	单选	A	一种特殊的机器学习算法	一种特殊的数据挖掘技术	一种特殊的数据可视化方法	一种特殊的数据清洗技术
304	数据预处理是指什么?	单选	D	将数据转换成特定的格式	删除无用的数据	检查和修复数据错误	所有以上答案

305	什么是特征选择?	单选	A	从大量特征中选择最相关的特征	从大量数据中选择最相关的数据	从大量模型中选择最相关的模型	从大量算法中选择最相关的算法
306	什么是聚类分析?	单选	A	将数据分成不同的类别的过程	将数据转换成特定的格式	存储数据的地方	数据分析的工具
307	什么是分类算法?	单选	A	一种用于将数据分成不同类别的算法	一种用于将数据转换成特定格式的算法	一种用于存储数据的算法	一种用于数据分析的算法
308	什么是回归分析?	单选	A	一种用于预测数值的分析方法	一种用于将数据转换成特定格式的分析方法	一种用于存储数据的分析方法	一种用于数据分析的分析方法
309	什么是关联规则挖掘?	单选	A	从大量数据中发现关联关系的过程	将数据转换成特定的格式	存储数据的地方	数据分析的工具
310	什么是异常检测?	单选	A	从大量数据中找出异常值的过程	将数据转换成特定的格式	存储数据的地方	数据分析的工具
311	什么是时间序列分析?	单选	A	一种用于分析时间相关数据的方法	一种用于将数据转换成特定格式的方法	一种用于存储时间相关数据的方法	一种用于数据分析的方法
312	什么是推荐系统?	单选	A	根据用户的兴趣和偏好推荐内容的系统	将数据转换成特定的格式的系统	存储数据的地方的系统	数据分析的工具的系统
313	什么是自然语言处理?	单选	A	计算机理解和处理人类语言的技术	将数据转换成特定的格式的技术	存储数据的地方的技术	数据分析的工具的技术
314	什么是数据工程?	单选	D	将数据转换成特定的格式的过程	将数据存储到数据库中的过程	将数据分析并提取有用信息的过程	所有以上答案
315	什么是数据科学?	单选	A	通过数据分析和解释来解决现实问题的学科	将数据转换成特定的格式的学科	存储数据的地方的学科	数据分析的工具的学科
316	什么是大数据?	单选	A	数据量超过常规处理软件能力的数据集	将数据转换成特定的格式的数据集	存储数据的地方的数据集	数据分析的工具的数据集
317	什么是云计算?	单选	A	通过网络提供计算和存储资源的方式	将数据转换成特定的格式的方式	存储数据的地方的方式	数据分析的工具的方式
318	什么是数据隐私?	单选	A	保护个人数据不被非法获取和滥用的措施	将数据转换成特定的格式的措施	存储数据的地方的措施	数据分析的工具的措施
319	什么是数据安全?	单选	A	保护数据不被非法获取、篡改和破坏的措施	将数据转换成特定的格式的措施	存储数据的地方的措施	数据分析的工具的措施
320	什么是数据治理?	单选	A	管理、保护和优化数据的过程	将数据转换成特定的格式的过程	存储数据的地方的过程	数据分析的工具的过程

321	什么是数据采集?	单选	A	收集数据的过程	将数据转换成特定的格式的过程	存储数据的地方的过程	数据分析的的工具的过程
322	什么是数据分析?	单选	A	通过分析数据来提取有用信息的过程	将数据转换成特定的格式的过程	存储数据的地方的过程	数据分析的的工具的过程
323	什么是数据可视化?	单选	A	通过图表和图形展示数据的过程	将数据转换成特定的格式的过程	存储数据的地方的过程	数据分析的的工具的过程
324	以下哪个不是AI的主要研究领域?	单选	D	机器学习	计算机视觉	数据挖掘	量子计算
325	机器学习是一种实现AI的方法, 它偏重于:	单选	B	利用规则和逻辑推理	使用大量数据和统计模型	模仿人类思维和行为	利用专家知识和经验
326	下列哪个是深度学习中常用的激活函数?	单选	A	sigmoid函数	二阶阶跃函数	正弦函数	指数函数
327	以下哪个算法是用于解决分类问题的?	单选	C	K-means算法	Apriori算法	决策树算法	PageRank算法
328	人工智能的发展历程中, 哪个阶段最早出现?	单选	B	强人工智能	弱人工智能	机器学习	深度学习
329	以下哪个是用于自然语言处理的技术?	单选	C	卷积神经网络	支持向量机	词袋模型	遗传算法
330	以下哪个是用于计算机视觉的技术?	单选	C	朴素贝叶斯	隐马尔可夫模型	卷积神经网络	遗传算法
331	以下哪个不是AI的应用领域?	单选	D	自动驾驶	金融风控	医学诊断	机器人制造
332	以下哪个算法是用于增强学习的?	单选	B	K-means算法K-means算法	Q-learning算法	PageRank算法	Apriori算法
333	对于监督学习任务, 以下哪个是正确的?	单选	C	训练数据集中不包含标签信息	模型可以自动学习规则和逻辑	训练数据集中包含输入和输出对应的标签信息	模型只能进行预测, 不能进行决策
334	以下哪个不是无监督学习的应用?	单选	D	聚类分析	异常检测	推荐系统	图像分类
335	以下哪个不是AI伦理问题?	单选	B	数据隐私	自我学习能力	就业岗位替代	偏见和歧视
336	以下哪个是AI系统的局限性?	单选	C	可以处理任何任务	没有容量限制	对数据的依赖性较强	不需要人类监督
337	以下哪个不是AI系统的优势?	单选	C	高速运算能力	学习和适应能力	无需电力供应	大规模数据处理能力
338	以下哪个不属于AI系统的核心技术?	单选	A	数据挖掘	自然语言处理	神经网络	机器视觉
339	神经网络中的感知机是由以下哪个科学家提出的?	单选	C	Marvin Minsky	John McCarthy	Frank Rosenblatt	Alan Turing
340	以下哪个不是AI系统的评估指标?	单选	D	准确率	召回率	F1分数	数据大小
341	以下哪个不是强化学习的关键元素?	单选	D	奖励信号	状态空间	动作空间	无监督学习
342	以下哪个不是AI系统的应用之一?	单选	D	语音识别	机器翻译	图像处理	数据存储
343	以下哪个不是AI系统的发展趋势?	单选	D	自主学习能力的提升	跨领域应用的拓展	人工智能与人类融合	取代人类智能
344	目前最先进的自然语言处理模型是哪个?	单选	A	BERT	LSTM	GAN	SVM

345	以下哪个是计算机视觉领域的经典数据集?	单选	D	CIFAR-10	MNIST	IMDB	ImageNet
346	以下哪个不是深度学习框架?	单选	D	TensorFlow	PyTorch	Keras	Scikit-learn
347	AlphaGo是哪家公司开发的?	单选	A	Google	Facebook	Microsoft	IBM
348	以下哪个不是人工智能的分支?	单选	C	机器学习	计算机视觉	自动控制	机器人技术
349	BAT是不是下面哪家企业的简称?	单选	A	京东	百度	腾讯	阿里巴巴
350	Python中的注释没有哪个符号() ?	单选	C	#	"" ""	//	""""
351	哪个不是智能交通可能会带来哪些好处()	单选	B	降低事故	提高速度	节能环保	缓解拥堵
352	下面哪个运算符不是逻辑运算符?()	单选	B	not	//	or	and
353	循环采集不包括()	单选	D	URL循环	文本循环	单个元素循环	随机循环
354	节点有多种类型,不包括是()	单选	D	元素、属性、文本	命名空间、处理指令	注释、文档节点	属性值
355	个人进行数据采集是,主要用于()	单选	A	网页信息收集	生产数据采集	市场数据采集	社会公开信息
356	处理噪声数据方法不包括()方法。	单选	D	分箱法	回归法	聚类	忽略元组
357	数据集成的方法不包括有()	单选	D	联邦数据库	中间件集成	数据仓库	聚类
358	数据规约的策略包括有()	单选	D	维归约	数量归约	数据压缩	实体识别
359	()是用电脑对文本集按照一定的标准进行自动分类标记。	单选	D	文本识别	机器翻译	问答系统	文本分类
360	人工智能对数据集进行数据标注时,以下符合标注规范的是()	单选	B	标注越多越好	标注框贴近目标	标注类别越多越好	标注命名可随意命名,不影响后期训练
361		单选					
362	以下不是数据标注员需要具备的素质	单选	D	学习力	责任感	专注力	有兴趣
363	关于python编程语言,下列描述正确的是()	单选	B	Python中整型有限制大小	append函数用于给列表增加元	del用于增加变量	Python中列表无法嵌套
364	下面关于函数说法错误的是()	单选	A	函数的三要素是函数名、参数和返回值,定义函数时这三个要素是必须的	如果没有return语句,则Python函数默认返回值为None	函数可以嵌套	使用def定义函数
365	Python的变量名可以由哪些元素组成()	单选	A	字母数字	运算符	符号	\$
366	哪个不是人工神经网络特点和优越性主要表现()。	单选	B	自学习功能	自动识别功能	高速寻找优化解的能力	联想存储功能
367	神经网络不可以按()分类	单选	C	学习方式分类	网络结构分类	网络的协议类型分类	网络的活动方式分类
368	专家系统的主要组成部分不包括	单选	D	知识库	推理引擎	用户接口	自主学习系统
369	机器学习相关算法不包括	单选	A	轨迹跟踪	决策树	数据挖掘	K近邻算法
370	以下不属于仿生算法的是()	单选	D	蚁群算法	遗传算法	人工神经网络	归并排序算法

371	关于卷积神经网络以下说法不正确的是	单选	B	常见池化层有最大池化与平均池化	卷积核不可以用来提取图片全局特性	处理图片时,是以扫描窗口的方式对图像做卷积	卷积神经网络可以包含卷积层,池化层和全连接层
372	以下是人工智能深度学习技术先寻找的是	单选	D	概率	数据	梯度	函数
373	下面对数据标注描述哪个是正确的()	单选	D	所有的数据标注都可以使用脚本语言自动标注	可以通过算法来实现数据的自动标注	OCR手写转录完全可以通过识别工具实现自动转录	数据标注是个重复性很强的工作
374	人工神经网络不包括	单选	C	输出层	中间隐藏层	映射层	输出层
375	数据标注前要对数据进行清洗以达到更好的训练识别效果,数据清洗方法下面说法不正确的是	单选	D	无效数据和缺失数据处理	数据一致性检查	数据定义	数据查重
376	下列不是图像处理时常用的?	单选	C	numpy	opencv	gensim	matplotlib
377	盲目搜索策略不包括下列那个	单选	D	广度优先搜索	深度优先搜索	有界深度优先搜索	全局择优搜索
378	人工智能中不是用“如果...则...”“关联起来的说法不包括	单选	B	产生式	规则	关系式	模式
379	下面不属于人工神经网络的是()	单选	C	卷积神经网络	循环神经网络	网络森林	深度信念网络
380	决策树中的分类结果是最末端的节点的有	单选	D	根节点	父节点	子节点	叶节点
381	机器学习从不同的角度,有不同的分类方式,以下哪项不属于按系统学习能力分类的类别	单选	D	监督学习	无监督学习	弱监督学习	函数学习
382	以下不属于人工智能在计算机视觉领域应用的是	单选	D	车站人脸识别进站	拍照识别植物	医疗影像诊断	实时字幕
383	人工神经网络的特点和优越性不包括_____	单选	B	自学习功能	自动识别功能	高速寻找优化解的能力	联想存储功能
384	不是专家系统组成部分的是	单选	A	深度学习	自然语言处理	专家系统	智能控制
385	如下能区分人说话的声音和小提琴的声音的是	单选	A	音质	音频	音量	音长
386	下列是直接影响传统机器学习算法成败的关键因素是哪个	单选	D	预处理	后处理	训练方法	特征提取
387	属于开区域标注的是	单选	D	曲线标注	多边形标注	框标注	线标注
388	关于python类 说法不正确的是	单选	B	类的实例方法必须创建对象后才可以调用	类的实例方法必须创建对象前才可以调用	类的类方法可以用对象和类名来调用	类的静态属性可以用类名和对象来调用
389	python有关异常说法正确的是	单选	B	程序中抛出异常终止程序	程序中抛出异常不一定终止程序	拼写错误会导致程序终止	缩进错误会导致程序终止
390	导入模块的方式不正确的是	单选	D	import mo	from mo import *	import mo as m	import m from mo

391	以下关于模块说法不正确的是	单选	C	一个xx.py就是一个模块	任何一个普通的xx.py文件可以作为模块导入	模块文件的扩展名不一定是.py	运行时会从指定的目录搜索导入的模块，如果没有，会报错异常
392	关于Python内存管理，下列说法不正确的是	单选	B	变量不必事先声明	变量无须先创建和赋值而直接使用	变量无须指定类型	可以使用del释放资源
393	知识图谱构建过程中不涉及下列哪些内容？	单选	D	知识获取	知识融合	知识验证	知识分析
394	下面哪项不属于神经网络的范畴？	单选	B	深度学习	机器学习	感知器	CNN
395	下面的Python开发包中，不属于深度学习框架的有	单选	A	flask	Tensorflow	Keras	Mxnet
396	数据采集的（）与技术执行的速度、团队内部成员协同以及数据分析需求和目标实现的效率有关。	单选	D	全面性	多维性	无序性	高效性
397	下列数据预处理任务，（）可以用来平滑数据，消除数据噪声。	单选	A	数据清洗	数据集成	数据变换	数据规约
398	下列不是用于大数据查询分析计算的产品是（）。	单选	B	Hive	HBase	SparkSQL	Dremel
399	数据仓库一般都是存在层次架构的，比如可以分为ODS、DW、DM层，不同粒度的表，不同作用的表，一般会分布在不同的层级中，如下相关信息表存放在ODS层的有（）。	单选	A	从源系统同步至数仓的粒度相同的一张登录日志表	日期为天粒度的登录统计表	从人的角度统计的登录情况表	日期为月粒度的登录统计表
400	从数据量方面来看，互联网系统和机器系统产生的数据量要（）企业系统的数据量。	单选	A	远远大于	远远小于	等于	无法确定
401	数据规约的目的是（）。	单选	C	填补数据中的空缺值	集成多个数据源的数据	得到数据集的压缩表示	规范化数据
402	下列陈述不的是（）。	单选	C	大数据将实现科学决策	大数据使政府决策更加精准化	大数据彻底将群体性事件化解在萌芽状态	大数据将实现预测式决策
403	下列哪项不是用于数据查询（）。	单选	B	Union	DELETE	SELECT	Project
404	AnalyticDB的优势不包括（）。	单选	D	更大规模和更快读写能力	更高可用和可靠性	更高安全	分布式计算能力

405	下列属于列族数据库的是（）。	单选	A	HBase	Redis	MySQL	MongoDB
406	Zookeeper的主要作用是（）。	单选	D	分布式的大量日志采集、聚合和传输	基于Hadoop的数据仓库	分布式列式数据库	分布式协调服务
407	利用Sqoop从关系数据库导入数据到Hive时，必须指定的参数为（）。	单选	C	hive-table	hive-database	hive-import	fields terminated by
408	下列不属于分布式计算技术的是（）。	单选	D	MapReduce	Spark	Flink	TensorFlow
409	下列哪项不是大数据发展的技术支撑（）。	单选	D	存储设备容量不断增加	网络带宽不断增加	CUP处理能力大幅提升	互联网数据量增加
410	数据合并（combine）是MapReduce Shuffle中一个重要环节，下列哪种应用不适合采用数据合并() class="markdown_return">。	单选	A	求平均值	求最大值	求最小值	求和
411	散点矩阵图用于哪种类型的数据（）。	单选	C	时间数据	比例数据	关系数据	文本数据
412	DataWorks底层分布式集群使用的是（）。	单选	C	Hadoop平台	Spark平台	飞天系统	天空系统
413	Spark的部署模式中哪种不是集群部署模式（）。	单选	A	本地模式	standalone模式	spark on yarn 模式	mesos模式
414	HBase数据库中，数据存储是按（）进行排序的。	单选	C	列族名称	列族限定符名称	行键	时间戳
415	（）被广泛用于购物篮分析。	单选	A	关联分析	分类分析	聚类分析	回归分析
416	关于大数据对人类思维的影响，不是的是（）。	单选	B	从“流程”核心转变为“数据”核心	由关注相关性转变为因果关系	从抽样转变为需要全部数据样本	从关注精确性转变为关注效率
417	（）将传统数据挖掘的思想和方法应用于Web，从Web资源和Web活动中爬取感兴趣的、潜在的、有用的模式和隐藏信息。	单选	A	Web挖掘	网络舆情	数据采集	离线浏览
418	关于探索性指标和报告性指标的描述错误的是（）。	单选	D	探索性指标是推测性质的，去发现一些未知的东西	报告性指标是关于公司日常运营、管理相关的指标	营业额属于报告性指标	销售量属于探索性指标

419	数据集成主要解决数据的分布性和（）问题。	单选	B	同构性	异构性	公开性	一致性
420	下列哪项不是Spark比MapReduce计算快的原因（）。	单选	D	基于内存的计算	基于DAG的调度框架	基于Lineage的容错机制	基于分布式计算的框架
421	关于HRegionServer功能描述，错误的是（）。	单选	D	响应用户I/O	向HDFS读写数据	HLog记录日志	实现Region负载均衡
422	从底层次数据抽象出高层次的描述过程叫做（）。	单选	C	抽样	离散化	数据立方体聚集	属性子集选择
423	下列关于MapReduce任务描述不的是（）。	单选	C	不同的Map任务之间不会进行通信	不同的Reduce任务之间不会发生任何信息交换	Map需要考虑数据全局性	用户不能显式地从一台机器向另一台机器发送消息
424	在一年一度的天猫双十一活动过程中，可视化大屏会动态实时展示销售额、订单量等指标，请问这些指标采用的是什么可视化图表（）。	单选	A	指标看板	仪表盘	极坐标	词云图
425	DataWorks的核心功能和特点不包括下列哪项（）。	单选	D	数据集成	数据开发	数据管理	数据分析
426	“全国疫情新增趋势”可能是可视化图形中的哪部分（）。	单选	D	坐标轴	刻度	图例	标题
427	通过大数据分析深入数据挖掘，无法实现（）。	单选	D	个性化营销和服务	实现预测性营销	洞察客户特性	提前消费
428	从数据产生速度来看，传统数据采集的数据几乎都是由人操作生成的，（）机器生成数据的效率。	单选	C	远远快于	等于	远远慢于	无法确定
429	下列（）是阿里云提供的底层的分布式计算平台。	单选	C	DataWorks	AnalyticDB	Maxcompute	DataV
430	第二次信息化浪潮的出现标志是（）。	单选	B	个人计算机开始普及	人类开始全面进入互联网时代	计算、大数据、物联网的快速发	人工智能技术高速发展

431	数据更重要的是能满足分析需求。灵活、快速自定义数据的多种属性和不同类型,从而满足不同的分析目标,这指的是数据采集的()。	单选	B	全面性	多维性	无序性	高效性
432	下列描述的是()。	单选	A	Hive支持分区	关系型数据不支持分区	Hive延迟低	关系型数据库延迟高
433	Hadoop生态的组件不包括()。	单选	D	HDFS	MapReduce	Hive	Flink
434	数据采集中的数据包括RFID数据、()数据、社交网络交互数据及移动互联网数据等海量数据。	单选	B	智能设备	传感器	温湿度	日志
435	大数据的多样性使得数据被分为三种数据结构,那么以下不是三种数据结构之一的是	单选	D	结构化数据	非结构化数据	半结构化数据	全结构化数据
436	下列选项中,不是人工智能的算法中的学习方法的是?	单选	A	重复学习	深度学习	迁移学习	对抗学习
437	自然语言处理难点目前有四大类,下列选项中不是其中之一的是	单选	A	机器性能	语言歧义性	知识依赖	语境
438	传统的机器学习方法包括监督学习、无监督学习和半监督学习,其中监督学习是学习给定标签的数据集。请问标签为离散的类型,称为分类,标签为连续的类型,称为什么?	单选	D	给定标签	离散	分类	回归
439	中国移动自主研发、发布的首个人工智能平台叫做()	单选	A	九天	OneNET	移娃	大云
440	HDFS中 Namenode的 Metadata的作用是什么?	单选	A	描述数据的存储位置等属性	存储数据	调度数据	数据
441	电信行业的客户关系管理中,客服中心优化可以实现严重问题及时预警,请问是用的是什么技术实现的?	单选	A	大数据技术	互联网技术	游戏技术	影像技术
442	随着闭源软件在数据分析领域的地盘不断缩小,老牌IT厂商正在改变商业模式,向着什么靠拢?	单选	B	闭源	开源	独立	封闭

443	以下不是非结构化数据的项是?	单选	C	图片	音频	数据库二维表数据	视频
444	以下数据单位换算错误的是?	单选	C	1KB=1024B	1GB=1024MB	1TB=1000GB	1MB=1024KB
445	下列选项中,不是Flume的特点的是?	单选	D	可靠性	集中式架构	可扩展性	可管理性
446	BP神经网络模型拓扑结构不包括	单选	D	输入层	隐层	输出层	显层
447	以下哪个不是语音识别的范畴?	单选	B	语音听写	语音台成	语音转写	语音唤醒
448	以下哪个场景可以称为大数据场景?	单选	C	故宫游客人	故宫门票收入	美团APP的定位信息	文章内容
449	2011年5月是哪家全球知名咨询公司在《Big data. The next frontier for innovation, competition and productivity》研究报告中指出,数据已经渗透到每一个行业和业务职能之中,逐渐成为重要的生产因素的?	单选	B	比尔恩门	麦肯锡	扎克伯格	乔图斯
450	下列选项中,哪项是分布式文件存储系统?	单选	A	HDFS	Flume	Kafka	Zookeeper
451	下列选项中,描述Flume对数据源的支持的是	单选	B	只能使用HDFS数据源	可以配置数据源	不能使用文件系统	不能使用目录方式
452	机器学习研究如何通过计算的手段,利用经验来改善系统自身的性能,请问机器学习利用数据训练出什么?	单选	A	模型	表结构	结果	报表
453	下列选项中,不是人工智能的算法中的学习方法的是	单选	A	重复学习	深度学习	迁移学习	对抗学习
454	语音识别产品体系有四部分,下列哪项不是体系之一?	单选	B	语音合成	语音播放	语音识别	语义理解
455	今年,大数据分析将出现革命性的新方法,从前的很多算法和基础理论可能会产生理论级别的突破。而哪项技术将继续成为大数据智能分析的核心技术	单选	A	机器学习	智能物流	脑科学	智能终端
456	以下哪个不属于大数据在电信行业的数据商业化方面的应用	单选	B	精准广告	网络管理	营销洞察	大数据检测和决策

457	RDD是由多个什么组成?	单选	A	partition	computer	Action	Transformation
458	下列选项中,不是大数据的一部分的是?	单选	D	海量计算	大量数据管理	数据分析	单机计算
459	属于“人造智能”,具有意识,达到或超越人类智慧水平的人工看能称为()	单选	C	高人工智能	低人工智能	强人工智能	弱人工智能
460	mapreduce计算模型适用于哪种任务?	单选	C	多线程处理	有关联的行处理	批处理	实时数据变化处理
461	当前世界产生的数据总量的单位是?	单选	B	KB	ZB	GB	TB
462	IBM提出的大数据5V特征包括更大(Volume)、更快(Velocity)、更多Variety)、更值钱(Value)和()	单选	C	更有效	更充分	更真实(Veracity)	更直观
463	Spark.是基于什么的迭代计算框架?它适用于需要多次操作特定数据集的应用场合。需要反复操作的次数越多,所需读取的数据量越大,受益越大,数据量小但是计算密集度较大的场合,受益就相对较小	单选	A	内存	硬盘	磁带	显卡
464	下列选项中,不是 kafka 适合的应用场景是?	单选	C	日志收集	消息系统	业务系统	流式处理
465	数据采集的基本步骤(ETL)中不包括哪项?	单选	C	抽取转换	加载	计算	
466	以下哪种学习方法不属于人工智能算法?	单选	D	迁移学习	对抗学习	强化学习	自由学习
467	语音识别常用的应用有四个,下列不是常用应用的是?	单选	D	聊天	拨号	导航	设备控制
468	下列选项中,哪项是由谷歌开发的人工智能算法框架?	单选	B	Kafka	Tensorflow	Caffe	Torch
469	“大数据”是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力来适应()、高增长率和多样化的信息资产。	单选	A	海量	少数	小量	结构化

470	MPP是指	单选	A	大规模并行处理系统	受限的分布式计算模型	集群计算资源管理框架	分布式计算编程框架
471	Hadoop是()年诞生的?	单选	C	1985-1985	1995-1996	2005-2006	2015-2016
472	在 Spark的软件栈中,用于流计算的是?	单选	C	GraphX	Mllib	Spark Streaming	SparkSQL
473	以下数据量可以称为大数据的是	单选	C	100MB	100KB	100PB	100MB
474	人工智能通过输入的图片,解析出图片的内容,这种技术叫什么?	单选	A	图片识别	语音识别	自动驾驶	消费金融
475	IBM的()是第一个在国际象棋上战胜人类 <br class="markdown_return">人工智能计算机。	单选	B	Alphago	深蓝	图灵机模型	深度学习机器人
476	下列选项中,不是人工智能的基础设施的是?	单选	D	CPU服务器	GPU服务器	专用芯片	游戏显卡
477	以下哪个不属于大数据在电信行业的应用?	单选	B	数据商业化	物流网络	企业运营管理	客户关系管理
478	客服中心是运营商和客户接触较为频繁的通道,我们可以利用()在后端给客服中心建立庞大的知识库,供接线员们使用帮忙他们快速精准的找到答案。	单选	A	大数据技术	客户画像	客户状态	客户心情
479	IBM提出的大数据5V特征包括()、更快(velocity)、更多(Variety)、更值钱(Value)和更真实(Veracity)	单选	B	更有效	更大(Volume)	更充分	更直观
480	Fume采用了三层架构,分别为 agent, collector和()	单选	B	Map	storage	Shuffle	Hash
481	Hadoop2系列版本中默认的HDFS的block是多大?	单选	C	32MB	64MB	128MB	16MB
482	工信部官网正式发布大数据产业“一三五”发展规划是在哪一年	单选	D	1987	1997	2007	2017
483	BP神经网络的学习规则是?	单选	B	梯度上升法	梯度下降法	梯度提升法	梯度曲线法

484	中国移动研发了智能客服问答机器人(),使业务流程自动化,更加智能的回答用户解决业务咨询、业务办理、流量查询等问题。	单选	D	精灵	阿尔法蛋	小度	移娃
485	对抗学习中两个网络互相竞争,一个负责生成样本,那么另一个负责做什么?	单选	A	判别样本	计算样本	统计样本	生成样本
486	以下应用没有使用你的地理位置信息的是?	单选	D	美团	滴滴	高德地图	word
487	下列选项中,不是大数据发展趋势的是	单选	D	大数据分析的革命性方法出现	大数据与云计算将深度融合	大数据一体机将持续发布	大数据未来可能会被淘汰
488	在 Spark的软件栈中,用于图计算的是	单选	C	Spark Streaming	Mllib	GraphX	SparkSQL
489	HBASE的特点不包括哪些?	单选	A	面向行	稀疏性	多版本	高可靠性
490	总体来说,人工智能发展的未来趋势是?	单选	A	上升	下降	不动	大幅度下降
491	以下数据单位从小到大排列的顺序是?	单选	B	GB、B、KB	B、KB、MB	KB、ZB、PB	B、MB、KB
492	最初的大数据概念还比较模糊,只是隐约的知道像个性化推荐、搜索引擎之类的处理需要大量数据,那么在搜索引擎方面,谁是世界上最大的厂商?	单选	A	谷歌	百度	360	bing
493	下列选项中,哪个不是 HBASE的特点?	单选	A	面向行	多版本	扩展性	稀疏性
494	下列选项中,不是用于数据存储的技术是哪一个?	单选	D	MongoDB	MySQL	HDFS	Java
495	下列不属于朴素贝叶斯算法优点的是 ()	单选	C	算法比较简单	对小规模的数据表现很好,能够处理多分类任务,适合增量式训练	对输入数据的表达形式很敏感	朴素贝叶斯模型发源于古典数学理论,有稳定的分类效率
496	检测一元正态分布中的离群点,属于异常检测中的基于 ()的离群点检测	单选	A	统计方法	邻近度	密度	聚类技术

497	通过对用户上网行为进行分析, 以实现精确智能营销, 可以使用()数据挖掘技术对客户进行分群	单选	D	分类	基于用户的协同过滤	关联分析	聚类
498	最近邻分类器的特点中, 不包括 ()。	单选	C	它使用具体的训练实例进行预测, 不必维护源自数据的模型	分类一个测试样例开销很大	最近邻分类器基于全局信息进行预测	可以生产任意形状的决策边界
499	将原始数据进行集成、变换、维度规约、数值规约是在 ()过程中的任务.	单选	C	频模式挖掘	分类和预测	数据预处理	数据流挖掘
500	在Python中, 下列关于集合说法错误的是 ()。	单选	A	集合具有互异性, 定义集合时允许出现相同的元素	集合具有确定性, 可以用in来判断元素是否在集合内	集合具有无序性, 集合没办法进行排序	集合具有运算性, 支持并交差等运算
501	下列属于深度神经网络模型的是	多选	ABCD	DNN深层神经网络	CNN卷积神经网络	RNN循环神经网络	GAN生成对抗网络
502	下列属于机器学习类型的是	多选	ABCD	有监督学习	无监督学习	半监督学习	强化学习
503	基因遗传算法的组成部分包括	多选	ABCD	初始化编码	适应度函数	选择	交叉和变异
504	基因遗传算法的两个常用的结束条件为	多选	AB	达到一定的迭代次数	适应度函数达到一定的要求	达到一定的变异次数	达到一定的交叉次数
505	一个完整的人工神经网络包括	多选	AC	一层输入层	多层分析层	多层隐藏层	两层输出层
506	在强化学习中, 主体和环境之间交互的要素有	多选	ABC	状态	动作	回报	强化
507	使用有监督学习的问题可以分为哪两大类	多选	AD	回归问题	抽样问题	聚类问题	分类问题
508		多选					
509	深度学习中以下那些步骤是由模型自动完成的	多选	AD	模型训练	特征选择	分析定位任务	特征提取
510	深度学习的主要过程包括 ()	多选	ABCD	选择适合问题的网络结构	选择适合网络结构的问题	用大量数据训练网络对权重初始化	优化网络
511	下列观点中, 属于符号主义的有	多选	ABCD	认为人的认知基元是符号	认为知识是信息的一种形式	认为人是一个物理符号系统	认为人工智能源于数理逻辑
512		多选					
513	常见的图像数据标注有哪几种类型	多选	ABCD	矩形框标注	关键点标注	区域标注	属性标注
514	为进行分类模型的训练和性能评价, 需要将输入的标注数据划分为	多选	BD	数据的类标	测试集	数据的特征	训练集
515	以下属于计算机视觉的是	多选	AC	物体识别和检测	语音导航	视觉问答	机器翻译
516	人工智能关键技术框架主要包括那两层	多选	BC	基础设施	算法	技术	人员
517	以下那些是人工智能的研究领域	多选	ABC	机器证明	模式识别	人工生命	变异原理

518	智能的特点是	多选	AC	能对环境进行灵活的应对	能够不断创新	具有十分牢固的记忆力	经济高效
519	机器翻译的局限性在于	多选	BC	训练样本单一	只能处理简单句	基于已有的既成案例	错误较多
520	机器智能的内涵包括	多选	ABCD	脑认知基础	机器感知与模式识别	自然语言处理和理解	知识工程
521	下面选项中()是人类特有,而机器所不具备的	多选	BCD	定量计算	规律总结	推理和直觉	广泛外延
522	设计一个自己的股票交易机器人需要做到	多选	ABCD	提出假设	建立模型	回测验证	执行交易
523	以下那些方面是机器人能够做到的	多选	ABC	医疗	围棋	写诗	灾害后救灾行为
524	机器智能种类包括	多选	BC	机器动作智能	机器行为智能	左右大脑加小脑功能	机器语言智能
525	自然智能包括	多选	CD	高级智能	超级智能	人类智能	机器智能
526	智慧教育的支撑技术包括	多选	ABCD	物联网	大数据	云计算	泛在网络
527	人工智能发展的驱动力包括	多选	ABCD	大数据	传感器	脑科学	超级计算
528	人工智能催生新技术、新模式、新业态、新产业、新产品,改变人类的	多选	ABCD	生产方式	生活方式	行为模式	思维模式
529	下列哪项属于自然智能	多选	ABC	植物	动物	细菌	机器
530	人工智能发展的三阶段是指哪三阶段	多选	ACD	推理期	思维期	知识期	机器学习期
531	下列属于感知智能的相关技术有哪几项	多选	ABCD	人脸识别	图像识别	语言识别	指纹识别
532	语音标注数据的典型应用场景包括	多选	ABCD	智能医疗	智能驾驶	智能家居	智能教育
533	不属于常见问题解答模块的主要技术的是	多选	BCD	问句相似度计算	语料库的构建	查询扩展	模式匹配
534	下列有关人工智能、机器学习、深度学习三者关系的说法正确的是	多选	CD	模式识别	文字合成	图像搜索	图像还原
535	不属于开区域标注的是	多选	A B C	曲线标注	多边形标注	框标注	线标注
536	下列选项中属于感知技术的是	多选	C D	机器学习技术	计算实施技术	语音识别技术	机器视觉
537	无人超市采用了()等多种智能技术,消费者在购物流程中将依次体验自动身份识别、自助导购服务、互动式营销、商品位置侦测、线上购物车清单自动生成和移动支付	多选	A B C D	计算机视觉	传感器定位	深度学习算法	图像分析
538	XML格式的文件,中哪一个是真的	多选	A C D	属性必须加引号	不区分大小写字母	有根元素	有关闭标签
539	属于闭区域标注的有	多选	A C	曲线标注	线标注	多边形标注	点标注
540	图像区域标注现在可以使用的手段有	多选	A C	手动标注	全自动化标注	半自动化标注	人工智能标注
541	智能眼镜可以看到对方说话内容的字幕,这个需要用的什么技术	多选	B C D	TTS	ASR	机器翻译	语音数据采集与标注
542	不能让计算机能够说话	多选	A C D	STT	TTS	ASR	OOT

543	按照语音信号处理研究方向语音数据标注任务可划分为?	多选	A B C D	语音识别	语音合成	情感识别	语音分离
544	语音编码格式是指按一定格式压缩采样和量化后的数值,从而降低音频的数据量,便于音频数据的存储和传输。常用的编码格式有	多选	B C	AVI	PCW (WAV)	MP3	JPG
545	语法类异常包括	多选	A B C	不规则的取值	值域格式错误	词法错误	数据中存在重复值
546	序列标注包括	多选	A B C	分词	意图理解	关键字	修饰关系
547	按照智能应用场景,语音数据标注任务可划分	多选	ABCD	智能家居	智能会议	智能客服	智能车载
548	ASR在中文领域有很大的难度,主要原因是	多选	AB	中文博大精深	方言众多	目前ASR技术不成熟	数据量不够
549	拉框标注通常用于自动驾驶应用中的	多选	ABD	行人标注	交通灯识别	车道线标注	汽车识别
550	关于信噪比错误的理解是	多选	ACD	信号与噪声之间的时间比	越高越好	对语音识别性能影响不大	一般真实的语音识别场景中,采样的音频信噪比都很高
551	常见的数据预处理方法有	多选	CD	数据挖掘	数据标注	数据清洗	信息脱敏
552	关于自然语言处理正确的是	多选	ABC	目前最先进的自然语言处理技术是基于深度学习模型的	自然语言处理技术达到人类智能的标准是通过图灵测试	文本标注的对象是自然语言文本	文本标注主要是用于无监督学习
553	以下是正确的Python字符串	多选	BD	'abc" ab"	'abc" ab'	"abc" ab"	"abc\" ab"
554	数据预处理是一种数据挖掘技术,包括	多选	ABC	数据清洗	数据集成	数据归约	数据标注
555		多选					
556	数据预处理是一种数据挖掘技术,包括	多选	A B C D	数据清洗	数据集成	数据归约	数据变换
557	计算机视觉的应用领域包括	多选	A B C D	医学影像	交通管理	工业制造	公安监控
558	以下应用场景中,()可以认为是计算机视觉的任务	多选	B C D	图像存储	图像/视频中文字的检测和识别	视频中感兴趣目标的检测、跟踪和定位	图像中前景物体与背景的分割
559	数字图像处理一般要完成的任务有	多选	A B C D	图像增强	图像变换	图像获取	图像识别
560	伴随着人工智能技术的发展,出现了多种新的交互方式,如	多选	AB C D	语音交互	情感交互	体感交互	脑机交互
561	AI=()	多选	B C D	终身学习	迁移学习	强化学习	深度学习
562	下列对人工智能芯片的表述,正确的是()	多选	A B C	一种专门用于处理人工智能应用中大量计算任务的芯片	能够更好地适应人工智能中大量矩阵运算	相对于传统的CPU处理器,智能芯片具有很好的并行计算性能	目前处于成熟高速发展阶段
563	人工智能的研究范畴广泛且复杂,其发展需要与()和社会科学等学科深度融合	多选	A B C D	计算机科学	数学	认知科学	神经科学

564	下列选项中，是人工智能的基础设施是	多选	A B C	CPU服务器	GPU服务器	智能芯片	游戏显卡
565	数据清洗是指	多选	ABC	包括检查数据一致性	发现并纠正数据文件中可识别错误的最后一道程序	包括处理无效值	包括数据标注
566	从技术角度看，人工智能的挑战包括	多选	A B D	能否保证人工智能的应用开发被用于正确的目标	智能系统开发时存在严重的缺陷，会产生不可预测的后果	人工智能的强大能力产生的负面效果可能是缓慢而大规模的	人工智能设计者在制作机器人时，会将自己的想法加入到机器人的思维系统中
567	语义标注中的自定义标签包括	多选	A B C	意图级别配置	功能配置	预识别配置	情感配置
568	目前外科手术领域的医用机器人的优点有	多选	AB	定位误差小	手术创口小	不需要人类医生进行操作	能够实时监控患者的情况
569	发展出图像识别成功率超越人类的人工智能的主要因素有	多选	AB	计算力的提升	大量数据驱动	社会关注度提升	人类专家规则的完善
570	显示不完整的处理方法，下列说法正确的是	多选	C D	不予处理	不管清不清晰全部框选属性+噪声	只显示一点点看不到题号、作答区、插图的，框选+噪声	可以看清题型的（题型属性+噪声），能看到的题号、作答区，插图都要框起来
571	框题的顺序按照我们的阅读顺序来，下列框题的顺序正确的是	多选	A D	插图和答题区：先框插图再框答题区	插图和答题区：先框答题区再框插图	表格和答题区：先框表格内答题区再框表格	表格和答题区：先框表格再框表格内答题区
572	数据仓库的基本特征包括	多选	A B C D	数据仓库的数据是面向主题的	数据仓库的数据是集成的	数据仓库的数据是稳定的	数据仓库的数据是随时间不断变化的
573	下列属于数据采集渠道的是	多选	A B C D	直接购买或共享行业数据	网络数据采集	自行采集	第三方合作委托采集
574	数据采集注意事项包括	多选	A C D	深度理解	根据个人认知	及时沟通	注意采集质量
575	认知语言更多的考虑	多选	C D	语法	词韵	语义	语用
576	以下框图顺序正确的是	多选	A C	从左到右，从上到下	从左到右，从下到上	大题-大题题号-小题-小题题号-答题区	大题-小题-大题题号-小题题号-答题区
577	下列为填空题题型的是	多选	A B C D	括号（）	下划线	圆圈○	方框
578	下列哪些是数据预处理的常用技术	多选	A B C D	数字属性的缺失值补0	LabelEncoder	one-hot encoder	CountVectorize
579	图像数字化需要经过的步骤包括	多选	A C	采样	裁剪	量化	旋转
580	数据清洗过程中，总归可以归为3个阶段，分别是	多选	B C D	数据加工	数据质量检查	数据校对	数据标准化
581	下列选项中，那两项可用于数据采集技术	多选	A C	Flume	Hive	Kafka	Mahout
582	以下用到语音识别技术的应用包括	多选	A C	苹果手机Siri	微信	百度地图	Word

583	大数据的业务应用处理需要经过那些流程	多选	A B C D	数据采集	数据清洗	数据建模	数据加工
584	自然语言处理的难点不包括以下那两项	多选	A C	语言独立性	语言歧义性	多国语言	语境
585	下面哪些用卷积神经网络处理效果比较好	多选	A B	人脸识别	手写数字识别	自然语言处理	机器翻译
586	下面哪些用循环神经网络处理效果比较好	多选	A B	语音识别	自然语言处理	图像识别	人脸识别
587	机器学习的核心要素包括	多选	A C D	数据	操作人员	算法	算力
588	下来哪些概念已被广泛应用于计算机视觉、自然语言处理、无人驾驶等领域?	多选	A B C D	机器学习	深度学习	强化学习	迁移学习
589	神经网络起源很早,但真正被大规模商用还是在本世纪,你觉得阻碍神经网络在上世纪发展的主要原因是什么	多选	C D	理论没有发展起来	科学界不重视	计算能力跟不上	标注数据不足
590	从机器学习预测目标数据的离散程度,可以将机器学习分类哪两类问题?	多选	A B	回归问题	分类问题	识别问题	判定问题
591		多选					
592	知识图谱在现代搜索引擎中有广泛应用,下面哪些可能是属于知识图谱构建过程中包含的内容?	多选	A B C D	实体及关系抽取	实体属性抽取	半结构化信息抽取	结构化数据融合
593	定义类如下: class hello():defshowInfo(sef):print(self.x)下面描述正确的是	多选	A C	该类不可以实例化	该类可以实例化	在pycharm 工具中会出现语法错误,说self没有定义	该类可以实例化,并且能正常通过对象调用showInfo()
594	Python可以应用以下那些领域	多选	A B C D	网站开发	人工智能	APP	数据科学
595	关于Python正确的是	多选	B C	Python中的异常只是系统抛出	使用try语句可以捕捉异常	用户可以自定义异常	异常不会终止程序的执行
596	生物特征识别技术包括	多选	A B C D	虹膜识别技术	指纹识别	DNA识别技术	声纹识别
597	生物特征识别技术是一门跨学科的技术,包括	多选	A B C D	计算机与光学	声学	生物传感器	生物统计学原理
598	智能控制的开发,目前认为有以下途径	多选	A C	基于数据挖掘的专家智能控制	基于遗传算法的软计算控制	基于人工神经网络的神经网络控制	以上说法都不对
599	联想存储的特点是	多选	ABCD	可以存储许多相关(激励, 响应)模式对	以分布、稳健的方式存储信息	即使输入激励模式完全失真时,仍然可以产生正确的响应模式	可在原存储中加入新的存储模式
600	智能制造系统有那些方面构成	多选	ABCD	智能制造	智能产品	智能生产	智能服务
601	选择以下关于人工智能概念的正确表述	多选	ABD	人工智能旨在创造智能机器该题无法得分	人工智能是研究和构建在给定环境下表现良好的智能体程序该题无法得分	人工智能将其定义为人类智能体的研究该题无法得分	人工智能是为了开发一类计算机使之能够完成通常由人类所能做的事该题无法得分
602	如下学科哪些是人工智能的基础?	多选	ABCD	经济学0.25	哲学0.25	心理学0.25	数学0.25

603	下列陈述中哪些是描述强AI(通用AI)的正确答案?	多选	AB	指的是一种机器,具有将智能应用于任何问题的能力0.50	是经过适当编程的具有正确输入和输出的计算机,因此有与人类同样判断力的头脑0.50	指的是一种机器,仅针对一个具体问题	其定义为无知觉的计算机智能,或专注于一个狭窄任务的AI
604	选择下列计算机系统中属于人工智能的实例	多选	AD	Web搜索引擎	超市条形码扫描器	声控电话菜单该题无法得分	智能个人助理该题无法得分
605	选择下列哪些是人工智能的研究领域	多选	ABC	人脸识别0.33	专家系统0.33	图像理解	分布式计算
606	考察人工智能(AI)的一些应用,去发现目前下列哪些任务可以通过AI来解决	多选	ABC	以竞技水平玩德州扑克游戏0.33	打一场像样的乒乓球比赛	在Web上购买一周的食品杂货0.33	在市场上购买一周的食品杂货
607	选择以下关于人工智能概念的正确表述	多选	ABD	人工智能旨在创造智能机器该题无法得分	人工智能是研究和构建在给定环境下表现良好的智能体程序该题无法得分	人工智能将其定义为人类智能体的研究该题无法得分	人工智能是为了开发一类计算机使之能够完成通常由人类所能做的事该题无法得分
608	如下学科哪些是人工智能的基础?	多选	ABCD	经济学0.25	哲学0.25	心理学0.25	数学0.25
609	下列陈述中哪些是描述强AI(通用AI)的正确答案?	多选	AB	指的是一种机器,具有将智能应用于任何问题的能力0.50	是经过适当编程的具有正确输入和输出的计算机,因此有与人类同样判断力的头脑0.50	指的是一种机器,仅针对一个具体问题	其定义为无知觉的计算机智能,或专注于一个狭窄任务的AI
610	选择下列计算机系统中属于人工智能的实例	多选	AD	Web搜索引擎	超市条形码扫描器	声控电话菜单该题无法得分	智能个人助理该题无法得分
611	选择下列哪些是人工智能的研究领域	多选	ABC	人脸识别0.33	专家系统0.33	图像理解	分布式计算
612	产品自有数据就是自身产品销售过程中产生的数据。	判断	正确				
613	调查问卷是以问题的形式系统的记载调查内容的一种印件。	判断	正确				
614	互联网数据分布在网页的不同位置,我们很难采集下来。	判断	错误				
615	Excel是按照数据结构来组织,存储和管理数据的仓库。	判断	错误				
616	商务数据指用户在电子商务网站购买商品的过程中,网站记录用户行为的大量数据。	判断	正确				
617	采集交易数据主要是为了通过数据分析评估客户价值,将潜在客户变为价值客户。	判断	正确				
618	评价数据主要以图片的形式出现。	判断	错误				

619	商务数据可以监控竞争对手的动态。	判断	正确				
620	商务数据不同帮助企业和个人共享客户信息。	判断	错误				
621	020数据主要有020电商平台数据和展销平台组成。	判断	正确				
622	数据采集又称数据获取，是利用设备或技术手段从现实环境及网络获取数据并放入系统内部进行使用。	判断	正确				
623	大多数互联网网页编写都是用HTML语言。	判断	正确				
624	常见的采集方法包括web爬虫采集和API接口采集。	判断	正确				
625	Web爬虫主要分为通用网络爬虫和聚焦网络爬虫。	判断	正确				
626	调用网站自身提供的应用程序编程接口，可以实现网络数据采集。	判断	正确				
627	数据采集工具分为编程类和可视化采集工具两类。	判断	正确				
628	数据采集工具可以针对某个主题从微博爬取相关信息。	判断	正确				
629	数据采集工具不能爬取学术信息。	判断	错误				
630	Python是一款服务器端解释性开源非编译脚本语言。	判断	正确				
631	数据采集器是进行数据采集的机器或者工具。	判断	正确				
632	数据采集器建立的任务也称之为规则。	判断	正确				
633	数据采集器采集不同字段时，建立一种规则就可以了。	判断	错误				
634	简易采集模式是利用系统内置模板进行数据采集的模式。	判断	正确				
635	简易采集可根据不同的参数进行不同程度的自定义采集。	判断	正确				
636	简易采集模板所有用户都可以使用。	判断	错误				
637	单页采集只能采集列表数据。	判断	错误				
638	表格数据采集要先选中表格中的一行数据。	判断	正确				
639	单网页采集主要是采集一个网页中的数据。	判断	正确				

640	列表详情页数据采集比单网页数据采集复杂。	判断	正确				
641	列表详情页需要从每个商品的标题进入到商品的详情页。	判断	正确				
642	采集列表详情页数据不需要循环步骤。	判断	错误				
643	URL循环是通过使用多个URL地址进行数据采集。	判断	正确				
644	分页循环采集是指把一页数据分成多页进行采集。	判断	错误				
645	网址中有多个页面需要采集叫分页循环采集。	判断	正确				
646	点击页面上的“下一页”按钮翻页，是最常见的翻页方式。	判断	正确				
647	需要登录时就不能进行采集。	判断	错误				
648	采集器无法识别验证码。	判断	错误				
649	Cookie登录的方式不需要输入账号和密码，直接打开网页就是登录状态。	判断	正确				
650	Cookie登录中Cookie长期存在，所以是登录采集中，最方便的采集方式。	判断	错误				
651	ajax可以通过在后台与服务器进行少量数据交换，可以重新加载整个网页。	判断	错误				
652	使用ajax技术的网页，点击网页中某个按钮或下拉页面，网址一般不发生改变，网址栏不出现加载状态，但网页局部有新的数据加载出来，有所变化。	判断	正确				
653	Ajax点击可以对一个按钮进行多次点击。	判断	错误				
654	如果设置了ajax技术的网页，在采集设置时，没有勾选ajax加载，则不能进行采集。	判断	错误				
655	XPath语言是网页内容定位语言，它可以帮助采集工具查找网页内容在网页中的位置。	判断	正确				
656	XPath语言，也称为HTML路径语言。	判断	错误				
657	节点关系是指节点与节点之间的关系，通过包含与被包含关系区分。	判断	正确				
658	XPath中，Text()函数，主要功能为选中指定文本内容的元素。	判断	正确				

659	、XPath中，contains(参数1,参数2)，主要功能为选中参数1 中包含参数2 中内容的元素。	判断	正确				
660	XPath中，last()，主要功能为选中同胞元素中最后一位的元素。	判断	正确				
661	XPath中，position()，主要功能为描述元素在同胞元素中的位置。	判断	正确				
662	使用采集器采集需要大量的专业知识。	判断	错误				
663	在公安经侦领域通过对网络敏感信息监控比对管理，帮助公安经济侦查工作的开展和提供数据层面的支持。	判断	正确				
664	商务数据采集多用于采购分析、市场分析、项目运行和论文编写。	判断	正确				
665	准确性是指数据是否正确的，数据存储在数据库中的值是否对应于真实世界的值。	判断	正确				
666	数据质量完整性是指信息具有一个实体描述的所有必需的部分	判断	正确				
667	空值是指缺失或不知道具体的值，可能是一条记录中的某个属性缺失，也可能是整条记录都丢失。	判断	正确				
668	数据质量的数据一致性是指在数据库中，不同表中存储和使用的同一数据应当是等价的，表示数据有相等的值和相同的	判断	正确				
669	数据质量的数据相关性是指数据与特定的应用和领域有关。	判断	正确				
670	数据质量的时效性是指有些数据会随时间而变化的	判断	正确				
671	数据质量的可信性由三个因素决定：数据来源的权威性、数据的规范性、数据产生的时间。	判断	正确				
672	数据质量的可解释性，也称为可读性，是指数据被人理解的难易程度	判断	正确				
673	由于操作员重复录入，并发处理等不规范的操作，导致产生不完整，不准确的，无效的数据也可以用在数据分析里。	判断	错误				
674	数据的缺失值是指现有数据集中某个或某些属性的值时不完整的。	判断	正确				

675	处理空缺值的基本方法有6种，包括忽略元组、人工填写空缺值、使用全局常量替换空缺值、使用属性的中心度量填充 空缺值、使用与给定元组属同一类的所有样本的平均值来填充空缺值、使用最可能的值填充空缺值。	判断	正确				
676	若一条记录中有多个属性值被遗漏了，则可将该记录排除在数据挖掘之外。	判断	正确				
677	噪声数据是指一个测量变量中的随机错误或偏差。	判断	正确				
678	噪声数据也可以用于数据挖掘进行分析。	判断	错误				
679	孤立点的是不符合数据模型的数据。	判断	正确				
680	孤立点并不真实存在，是人们无意中弄出的偏差比较大的数据。	判断	错误				
681	数据集成是指将互相关联的分布式异构数据源集成到一起，使用户能够以透明的方式访问这些数据源。	判断	正确				
682	数据仓库是数据集成最常用的方法。	判断	正确				
683	实体识别，它所解决的问题是如何匹配多个信息源在现实世界中的实体事物。	判断	正确				
684	数据规约就是为了压缩数据量，帮助从原有庞大数据集中获得一个精简的数据集合，并使这一精简数据集保持原有数据 集的完整性。	判断	正确				
685	主成分分析也称主分量分析，旨在利用降维的思想，把多指标转化为少数几个综合指标。	判断	正确				
686	文本标注需要按照自然语言处理的要求进行标注，其中自然语言处理的英简称是NLP。	判断	正确				
687	情感语音方面的研究理论与方法尚不成熟，而人类情感具有复杂性和个性化的特点，给其研究带来诸多困难。	判断	正确				
688	语音转换可以理解为如下语音处理过程：先通过语音识别将输入语音转换为文字，对文字进行转换，再将转换之后的文字通过语音合成技术输出新的语音。	判断	错误				
689	人工智能项目中，标注环节没必要重视。	判断	错误				

690	一个合格的标注员需要具备以下素质：持续的学习力、细心、耐心、责任心、较强的专注力、团队协作、良好的沟通表达能力等	判断	正确				
691	人工智能算法的训练一般需要训练集、测试集和验证集。	判断	正确				
692	计算机具有从图像中识别物体的能力，但是图像噪音较大时识别率不高。	判断	正确				
693	自然语言是人类创造，是一种为某些特定目的而创造的语言。	判断	错误				
694	尺寸标注表示的是机件的真实大小，应以图样上所注的尺寸数值为依据，与图形的大小及绘图准确度无关	判断	正确				
695	线性尺寸的数值一般注写在尺寸线的下方或中断处。	判断	错误				
696	学习率越大，训练速度越快，最优解越精确。	判断	错误				
697	计算机视觉的任务还包括人眼不擅长的工作，例如图像中涉及场景、人物的三维重建。与很多学科都有密切关系，例如数字图像处理、模式识别、机器学习、计算机图形学等	判断	正确				
698	为提升效率，大多数AI框架的核心功能模块都是C++实现的	判断	正确				
699	机器人一般由执行机构、驱动装置、检测装置、控制系统和复杂机械等组成，涉及到控制论、机械电子、计算机、材料、仿生等学科，在工业、医学、农业、建筑业甚至军事等领域中均有重要用途	判断	正确				
700	强人工智能观点认为可能创造出真正推理和解决问题的智能机器。	判断	正确				
701	机器的优势在于善于处理复杂的确定问题。	判断	正确				
702	人工智能读片的方法之一是利用目标检测。	判断	正确				
703	智能家居应该自动感知周围的环境，可以不需要人的操控。	判断	正确				

704	智能音箱本质上是音箱、智能语音交互系统、互联网、内容叠加的产物。	判断	正确				
705	根据发展趋势定义,人工智能就是会不断自我学习的计算机程序。	判断	正确				
706	人工智能学习玩Flappy Bird过程中,只需要人类告诉AI不能碰到水管即可,不需要提供其他信息。	判断	错误				
707	医疗健康领域,人工智能在医学影像方面的应用被认为最不可能率先实现商业化。	判断	错误				
708	只有符合社会伦理规范和公共政策的解决方案,才能设计出可信赖的人工智能。	判断	正确				
709	重复性强,要求弱社交能力的工作是最容易被AI取代的工作。	判断	正确				
710	强人工智能观点认为可能创造出真正推理和解决问题的智能机器。	判断	正确				
711	尺寸界线用细实线绘制,并由图形的轮廓线、轴线或对称中心线处引出,不可直接以这些线作为尺寸界线。	判断	错误				
712	依据《人工智能训练师国家职业技能标准》命题,学生组对应那个技能等级技师	判断	错误				
713	自然语言处理并不是一般地研究自然语言,而在于研制能有效地实现自然语言通信的计算机系统,特别是其中的软件系统,但它并不属于计算机科学的一部分。	判断	错误				
714	Json格式的数据就是python中的字典。	判断	错误				
715	Python中的集合数据类型中的元素是有序的。	判断	错误				
716	Python是一种跨平台、开源、免费的高级动态编程语言。	判断	正确				
717	Python中的集合数据类型中的元素是有序的。	判断	错误				
718	Python中的变量在使用前必须先定义。	判断	错误				
719	已知 $x = 3$, 那么赋值语句 $x = 'abcdefg'$ 是无法正常执行的。	判断	错误				
720	$3+4j$ 不是合法的Python表达式。	判断	错误				
721	9999**9999这样的命令在Python中无法运行。	判断	错误				

722	Python代码的注释只有一种方式，那就是使用#符号。	判断	错误				
723	Python支持使用字典的“键”作为下标来访问字典中的值。	判断	正确				
724	列表可以作为字典的“键”。	判断	错误				
725	、已知x为非空列表，那么表达式sorted(x, reverse=True) == list(reversed(x)) 的值一定是True。	判断	错误				
726	生成器推导式比列表推导式具有更高的效率，推荐使用。	判断	正确				
727	Python列表、元组、字符串都属于有序序列。	判断	正确				
728	在Python 3.x中语句 print(*[1,2,3]) 不能正确执行。	判断	错误				
729	列表对象的append()方法属于原地操作，用于在列表尾部追加一个元素。	判断	正确				
730	在UTF-8编码中一个汉字需要占用3个字节。	判断	正确				
731	如果仅仅是用于控制循环次数，那么使用for i in range(20)和for i in range(20,40)的作用是等价的。	判断	正确				
732	机器学习的学习方式是通过获得经验或历史数据不断改进提高做某项任务的表现。	判断	正确				
733	机器学习是一门多领域交叉学科，涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。	判断	正确				
734	python中如果导入cv2，如下python代码cap=cv2.VideoCapture('video.mp4')ret,frame = cap.read()返回值ret是布尔型，正确读取则返回True，读取失败或读取视频结尾则会返回False，frame为最后一帧的图像	判断	错误				
735	python中如果导入cv2，使用cap = cv2.VideoCapture(0)表示打开笔记本的内置摄像头	判断	正确				

736	计算机视觉(Computer Vision)是指利用摄像机和电脑代替人眼,使得计算机拥有类似于人类的那种对目标进行分割、分类、识别、跟踪、判别决策的功能。	判断	正确				
737	cd到你的python解释器下的scripts中,可以使用pip install + (库的名称) 安装库	判断	正确				
738	使用pip命令安装库的时候,为了提高效率,有时会使用国内镜像安装库,命令格式是pip install + 库名 + -m + 镜像地址	判断	错误				
739	from A import B是从模块库B中导入A模块	判断	错误				
740	import sys as s是导入库s,别名是sys	判断	错误				
741	所有的机器学习分类算法都属于有监督方法。	判断	正确				
742	监督学习的数据必须要带标签等人为标注信息。	判断	正确				
743	pytorch通常需要用户编写自定义训练循环,训练循环的代码风格因人而异。有3类典型的训练循环代码风格:脚本形式训练循环,函数形式训练循环,类形式训练循环。	判断	正确				
744	对抗神经网络可以通过两个神经网络的博弈,达到更好的学习效果。 正确	判断	正确				
745	监督学习需要有大量标注好的数据集来对模型进行训练。	判断	正确				
746	传统的机器学习方法的表现主要算法,目前的机器学习主要是强化学习,具有自学习能力。	判断	正确				
747	算法做出的价格往往比人做出的市场价格要合理。	判断	错误				
748	人工智能算法是严格按照规则进行股市交易的。	判断	正确				
749	RNN循环神经网络是用来处理文本、视频、音频等序列数据。	判断	正确				
750	DNN是卷积神经网络。	判断	错误				

751	深度学习是基于RNN循环神经网络。	判断	错误				
752	监督学习是从标记的训练数据来推断一个功能的机器学习任务。	判断	正确				
753	根据类别未知(没有被标记)的训练样本解决模式识别中的各种问题,称之为监督学习。	判断	错误				
754	根据类别未知(没有被标记)的训练样本解决模式识别中的各种问题,称之为无监督学习。	判断	正确				
755	启发式算法与AlphaBeta剪枝类似,是从叶节点自底向上计算估值。	判断	错误				
756	仿生算法是一类模拟自然生物进化或者群体社会行为的随机搜索方法的统称。	判断	正确				
757	在解决函数优化问题时,基因遗传算法的全局性不好,容易陷入局部最优值。	判断	错误				
758	自然界中生物变异的概率是不确定的,但是基因遗传算法的变异概率可以人为调节。	判断	正确				
759	神经网络中各个隐藏层能提取出和人类看到的一样的特征。	判断	错误				
760	<pre>from PIL import Image import matplotlib.pyplot as plt img = Image.open(os.path.join('images', '2007_000648' + '.jpg')) plt.figure(figsize=(10,5)) #设置窗口大小 plt.suptitle('Multi_Image') # 图片名称 plt.subplot(2,3,1), plt.title('image') plt.imshow(img) </pre> <p>如上代码中plt.subplot(2,3,1)的意思是在同一画面创建3行2列个图形位置,当前图片现在在第1个位置</p>	判断	错误				
761	梯度下降算法是最常用也是最有效的神经网络的优化办法,完全可以满足不同类型的需求。	判断	错误				

762	Tensorflow是当下最流行的深度学习框架之一。	判断	正确				
763	Tensorflow是一个用于机器学习和深度学习的端到端开源平台。	判断	正确				
764	朴素贝叶算法不需要样本特征之间的独立同分布。	判断	错误				
765	循环神经网络所使用的训练法则叫做时序反向传播，简称BPTT。	判断	正确				
766	<pre>from PIL import Image img = Image.open(os.path.join('images', '2007_000648' + '.jpg')) gray = img.convert('1')</pre> <p>最后一句代码是将图片转化为一值图像</p>	判断	正确				
767	机器学习的含义是指机器面对自行为的修正或性能的改善和机器对客观规律和发展。	判断	正确				
768	深度学习在人工智能领域表现并不突出。	判断	错误				
769	监督学习的学习数据既有特征(feature)，也有标签(label)。	判断	正确				
770	相对于人工神经网络和深度学习，类脑人工智能对人类大脑的神经回路具有更深入的理解。	判断	正确				
771	类脑人工智能是指模拟人类大脑的人工智能。	判断	错误				
772	类脑人工智能及人工神经网络只是智能的一种形式。	判断	正确				
773	python做矩阵操作一般要先引入相应的库，例如 <code>import numpy as np</code>	判断	正确				
774	<pre>x=np.array([0, 1, 2, 3, 4]) y=x[::-1]</pre> <p>要将x和y相乘，如果np是导入的numpy库，使用的语句是 <code>np.dot(x, y)</code></p>	判断	正确				
775	机器学习算法训练完成后，在训练集准确率很高，但验证集准确率很低，说明已经过拟合	判断	正确				

776	专家系统模拟人类专家的知识和经验解决特定领域的问题，实现了人工智能从理论研究走向实际应用。	判断	正确				
777	对于自然语言处理工程来说，良好的语料标注是项目成功的基础。	判断	正确				
778	可以结合人工分词来加快文本数据标注的进度。	判断	错误				
779	语料标注涵盖在自然语言处理任务的建模和标注两个步骤中，语料标注过程尽量严格按照标注规范可以确保一次性建立完美的标注模型。	判断	错误				
780	对专业度要求比较高的标注，最好请制定标注标准的人参与标注。	判断	错误				
781	区域标注指的是将图像分成各具特性的区域并提取出感兴趣部分的过程。	判断	正确				
782	“回归问题和分类问题都有可能发生过拟合”，这句话是否正确？（）	判断	正确				
783	对于回归问题和分类问题，最常用的指标都是准确率和召回率吗？（）	判断	错误				
784	CNN的全称是卷积神经网络，是否正确？（）	判断	正确				
785	RNN的全称是卷积神经网络，是否正确？（）	判断	错误				
786	MLP的全称是指卷积神经网络，是否正确？（）	判断	错误				
787	GAN的全称是指卷积神经网络，是否正确？（）	判断	错误				
788	给定n个数据点，如果其中一半用于训练，另一半用于测试，则训练误差和测试误差之间的差别会随着n的增加而减少，请问这句话是否正确。（）	判断	正确				
789	一般来说，回归不能用于分类问题，但是Logistic回归可用于解决分类问题，是否正确？（）	判断	正确				
790	训练CNN时，可以对输入进行旋转、平移、缩放等预处理提高模型泛化能力。这种说法是否正确？（）	判断	正确				

791	循环神经网络属于一种特殊的递归结构，这句话是否正确（）	判断	正确				
792	淘宝中我们输入关键词查看商品时，会自动搜索出很多和所搜商品相关联的信息，这就是利用知识图谱的典型用例。	判断	正确				
793	感知机的运作原理是逻辑判断流程。	判断	错误				
794	语音增强的主要任务就是消除环境噪声对语音的影响。	判断	正确				
795	语音识别中，最简单的是特定人、小词汇量、孤立词的语音识别。	判断	正确				
796	有更多隐层神经网络模型，我们称之为深度学习。	判断	正确				
797	神经元受到足够强度的刺激，才会响应释放出刺激其他神经元的递质，刺激不足不会有输出。	判断	正确				
798	有更多隐层神经网络模型，我们称之为深度学习。	判断	正确				
799	在图像处理中，采样越细，像素越小，越能精确的表现图像。	判断	正确				
800	判断：Python是通用编译器将程序编译成二进制代码再运行。	判断	错误				
801	感知机的运作原理是逻辑判断流程。	判断	错误				
802	一个实体指向它的属性值。不同属性类型对应于不同类型属性的边。（）	判断	正确				
803	常见的非结构化数据主要是文本类的文章，即自然语言数据。	判断	正确				
804	学习率越大，训练速度越快，最优解越精确。	判断	错误				
805	国家、民族、书籍、电脑等都是由实体组成的集合，即语义类(概念)	判断	正确				
806	不同于基于关键词搜索的传统搜索引擎，知识图谱可用来更好地查询复杂的关联信息，从语义层面理解用户意图，改进搜索质量。（）	判断	正确				
807	逆文档频率值越大，重要性越大。	判断	正确				
808	某个人、某个城市、某种植物在知识图谱中都被称为实体。	判断	正确				

809	目前出现的基于搜索引擎、问答系统的产品有百度“贴心”、搜狗“知立方”等。	判断	正确				
810	语音识别技术中最复杂最难解决的是，特定人、大词汇量、连续语音识别。	判断	错误				
811	聚类需要从没有标签的一组输入向量中寻找数据的模型和规律。	判断	正确				
812	监督学习的学习数据既有特征(feature)，也有标签(label)。	判断	正确				
813	学习率越小，训练速度越慢，最优解越精确。	判断	正确				
814	计算机视觉利用二维投影图像来重构三维物体的可视部分。	判断	正确				
815	人耳的精妙结构决定了我们对不同频率的声音有着不同的敏感。	判断	正确				
816	在梯度分别为正负的两个点之间，一定存在一个梯度为0的点。	判断	正确				
817	集合中的元素没有特定顺序但可以重复。	判断	错误				
818	聚类和分类的区别在于用于聚类的训练样本的类标记是未知的。	判断	正确				
819	机器学习算法在图像识别领域的性能表现可能会超过人类。	判断	正确				
820	列表是不可变对象，支持在原处修改。	判断	错误				
821	决策树学习是一种逼近离散值目标函数的方法，学习到的函数被表示为一棵决策树。	判断	正确				
822	梯度下降，就是沿着函数的梯度（导数）方向更新自变量，使得函数的取值越来越小，直至达到全局最小或者局部最小。	判断	正确				
823	学习率越大，训练速度越快，最优解越精确。	判断	错误				
824	线性回归是一种有监督机器学习算法，它使用真实的标签进行训练。	判断	正确				
825	最小二乘法是基于预测值和真实值的均方差最小化的方法来估计线性回归学习器的参数 w 和 b 。	判断	正确				
826	精确率、查全率、准确率、F1分数四种指标各有侧重，指标值都是越高越好，最佳值均为1，最差值均为0。	判断	正确				

827	查全率越高，意味着模型漏掉的样本越少，当假阴性的成本很高时，查全率指标有助于衡量模型的好坏。	判断	正确				
828	列表、元组和字符串都支持双向索引，有效索引的范围为 $[-L, L]$ ， L 为列表、元组或字符串的长度。	判断	错误				
829	列表、元组和字符串属于有序序列，其中的元素有严格的先后顺序。	判断	正确				
830	一般的，一棵决策树包含一个根结点、若干个内部结点和若干个叶结点；叶结点对应于决策结果，其他每个结点则对应于一个属性测试，根结点包含样本全集。	判断	正确				
831	Sigmoid函数能够把输入的连续实值变换为0和1之间的输出。	判断	正确				
832	监督学习的学习数据既有特征(feature)，也有标签(label)。	判断	正确				
833	线性回归主要用于解决回归问题，其因变量是连续的值。	判断	正确				
834	k均值算法，是一种原型聚类算法。	判断	正确				
835	聚类生成的组称为簇，簇内任意对象之间具有较高的相似度，而簇间任意对象之间具有较高的相异度。	判断	正确				
836	训练集与验证集的样本是不同的。	判断	正确				
837	在各类机器学习算法中，过拟合和欠拟合都是可以彻底避免的。	判断	错误				
838	寻找最优超参数费时费力，应该在模型训练之前就指定最优参数。	判断	错误				
839	准确率是所有正确识别的样本占样本总量的比例。当所有类别都同等重要时，采用准确率最为简单直观。	判断	正确				
840	Pandas中利用merge函数合并数据表时默认的是内连接方式。	判断	正确				
841	同一个列表中的元素的数据类型可以各不相同。	判断	正确				
842	回归中利用最小二乘法主要通过最小化误差的平方来寻找一个数据匹配的最佳函数。	判断	正确				
843	均方根误差与均方误差二者没有关系。	判断	错误				

844	在MapReduce计算模型中，数据经过Reduce阶段处理后传送给Map阶段继续处理。	判断	错误				
845	数据库是可共享、具有独立性的数据集合。	判断	正确				
846	Hive用于对数据进行分析，本身不存储数据，数据存储在HDFS系统中。	判断	正确				
847	Spark基于硬盘计算的大数据并行计算框架，可用于构建大型的、低延迟的数据分析应用程序。	判断	错误				
848	Hadoop的可扩展性错，不能随意增加节点。	判断	错误				
849	Nutch项目开发了自己的分布式文件系统NDFS。	判断	正确				
850	关系数据库可以存储结构化、非结构化数据。	判断	错误				
851	Hadoop主要用于存储结构化数据。	判断	错误				
852	DataWorks是阿里云提供的一站式大数据开发处理平台。	判断	正确				
853	分布式文件系统的设计目标是高性能、可伸缩性、可靠性及高安全性。	判断	错误				
854	Flink的流水线运行时系统可以执行批处理和流处理程序，但是不支持迭代算法。	判断	错误				
855	商业领域产生的数据量非常庞大，消费者的行为都会成为对企业非常有价值的数据来源。	判断	正确				
856	传统数据采集的数据都是非结构化的数据。	判断	错误				
857	HBase数据库中，用户读写操作通过HMaster实现。	判断	错误				

858	在噪声数据中，波动数据比离群点数据偏离整体水平更大。	判断	错误				
859	Hadoop分布式并行计算框架的执行是以任务展开，Storm分布式并行计算框架任务则是以提交拓扑的方式开始。	判断	正确				
860	异常数据都是需要删除处理的。	判断	错误				
861	HDFS是Hadoop架构中提供的数据存储组件。	判断	正确				