

# 考试：人工智能伦理 (COMP90087\_2021\_SM1)

开始时间：6 月 24 日 15:00

## 测验说明

### 学术诚信宣言

通过开始和/或提交此评估，我同意我已阅读并理解[大学关于学术诚信的政策。](https://academicintegrity.unimelb.edu.au/#online-exams)

(<https://academicintegrity.unimelb.edu.au/#online-exams>)

我也同意：

1. 除非第 2 款适用，否则我提交的作品将是原创的，并且完全是我自己的作品（作弊）；
2. 我不会寻求或接受任何其他人（串通）的任何帮助，除非工作是为了指定的协作任务，在这种情况下，将注明个人贡献；和，
3. 未经适当承认或引用（抄袭），我不会使用任何来源。
4. 如果我提交的作品是计算机程序或代码，我将确保：
  - a. 我复制的任何代码都通过在程序开头或头文件中标识该代码的来源来清楚地注明，或者，内联注释标识了复制代码的开头和结尾；和
  - b. 对来自其他地方的代码的任何修改都将被评论以显示修改的性质。

## 简答题 [4 分]

### 问题 1

0.5分

考虑以下场景，一名在诊所工作的医学专家诊断常见的皮肤问题，其中一些可能是严重的医疗问题。诊所有一个基于软件的系统，可以自动诊断皮肤问题，专家可以对照自己的诊断结果来检查自己的诊断，并可以探索预测的原因。

将以下案例与使用 Parasuraman 和 Riley 信任框架的使用类型相匹配。

诊所决定只使用自动诊断系统，因为它可以更快地返回结果并且与医学专家一样准确

滥用自动化



专家接受对常见问题的诊断，因为他们从未见过系统错误地预测了这种情况

有保证的信任



专家拒绝系统的良好诊断，因为他们有时不同意系统

有根据的不信任



专家忽略系统的正确诊断，因为专家认为他们做出了更好的诊断

不使用自动化



## 问题2

0.5分

关怀伦理是一种起源于 1970 年代和 1980 年代的道德理论。该理论源于（选择一个或多个答案）：

☒ 女权主义伦理

☐ 文化相对主义

☐ 科技研究

☐ 应用伦理

☐ 道义伦理

## 问题 3

0.5分

原则主义将理论的一些核心特征提炼为伦理原则，例如功利主义、美德伦理、义务论和伦理或关怀。下列哪一项**不是**原则主义的原则？（选择一个或多个答案）：

☒ 仁

☒ 恶作剧

☐ 尊重自主权

☐ 正义

☐ 行善

☐ 不要伤害**问题 4****0.5分**

在部署基于人工智能的招聘系统时，以下哪一项是减少潜在不平等的方法？

(A) 在入围候选人时应用鲁尼规则和其他平权政策。

(B) 仅在系统用于“生产”的第一年审核系统的结果。

☒ (A) 仅。

☐ (B) 仅。

☐ (A 和 B) 两者。

☐ (A或B) 都没有。

**问题 5****0.5分**

当用户安装 Windows 10 时，他们会看到一个选项来确定要向 Microsoft 发送多少诊断信息。

诊断信息“用于帮助保持 Windows 安全和最新，解决问题并改进产品”  
(Microsoft, 2021) 。

但是，用户只能选择“基本”诊断数据或“完整”诊断数据；并且没有选择退出或禁用该功能的选项。为了让事情变得混乱，还有另一种选择.....

“让 Microsoft 根据诊断数据为您提供量身定制的体验.....”包括“个性化提示、广告和建议” (Microsoft, 2021 年) 。

这说明了 \_\_\_\_ 透明度的现象，其中用户被给予了最低水平的透明度，这并不能完全支持他们的决策。

☐ 可解释的

☐ 算法的

☒ 名义上的

☐ 有效的

## 问题 6

0.5分

一家大型超市使用机器学习来确定每周每种产品所需的库存量。他们使用线性回归和只有三个属性来训练。在每一次每周例会上，超市经理都会得到一份建议，并解释为什么要提供这个建议。解释是直接从线性方程的权重中提取的。这给出了哪些类型的可解释性？

- ☐ 本地、事后可解释性
- ☐ 本地化、可解释性
- ☐ 全局、事后可解释性
- ☐ 全球性、可解释性

## 问题 7

0.5分

Beard 和 Longstaff 框架中的非工具主义原则需要一种技术来对待人

[ 选择 ]



不仅仅是作为

[ 选择 ]



. 这

一原则的基础是人

[ 选择 ]



.

## 问题 8

0.5分

根据 ICCPR 和 ICESCR 等国际人权条约，个人可以针对：

- ☐ 他们自己的政府，如果它已经签署了条约
- ☐ 商业
- ☐ 所有政府和企业

☐ 其他公民

# 书面回答问题 [16 分]

## 问题 9





4分

您所在的组织正在开发机器学习模型以预测审计项目的完成时间。这些模型具有 80% 的预测完成日期的准确度，并且向潜在客户宣传该准确度度量。

您的组织发现对应用程序的信任度很低。您的一位同事建议使用不同的准确度衡量标准：报告准确度为 90% 而不是 80% 的衡量标准。您同意这两种衡量方式都是衡量准确性的同样有效的方法，并且没有科学理由选择一种而不是另一种。

用 3-4 句话描述您是否相信这会与新客户建立信任。

编辑 看法 插 格式 工具 桌子

12pt ▾ 段落 ▾ | **B** *I* U A ▾  ▾ T<sup>2</sup> ▾ |  ▾  ▾  ▾ |





⋮

问题 10





4分

政府计划引入一个新的人工智能系统来确定国家残疾保险计划 (NDIS) 的付款。NDIS 旨在为残障人士提供资源。该系统将帮助决定谁有资格获得付款并分配适合每个人需要的金额。简要解释政府可能对这个人人工智能系统**负责的**四种方式。

编辑 看法 插 格式 工具 桌子

12pt ▾ 段落 ▾ | **B** *I* U A ▾  ▾ T<sup>2</sup> ▾ |  ▾  ▾  ▾ |

⋮

磷   | 0 字 |   ⋮

问题 11

4分

一家销售 Doge 品牌商品的电子商务公司 DogeToMoon 使用手动处理来在线完成客户订单。他们将客户数据存储在销售办公室桌面上的纯文本文件中。此外，销售人员不会验证数据，并且经常在客户姓名和地址中输入错误。

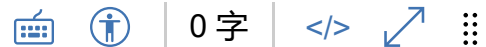
命名两个最相关的GDPR（一般数据保护条例）的数据保护原则DogeToMoon违反。对于 EACH 原则，请用简短的句子解释它是如何应用的。

编辑 看法 插 格式 工具 桌子

12pt ▾ 段落 ▾ | **B** *I* U A ▾  ▾ T<sup>2</sup> ▾ |  ▾  ▾  ▾ |

⋮

磷

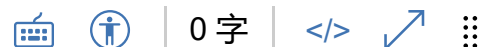
**问题 12****4分**

用三到四句话提供一个论据，哪个是解决人工智能对社会影响的更好方法：人权还是道德推理。确保你认为哪个更好。

编辑 看法 插 格式 工具 桌子

12pt ▾ 段落 ▾ | **B** *I* U A ▾  ▾ T<sup>2</sup> ▾ |  ▾  ▾  ▾ |  
⋮

磷



Saved at 13:13

提交测验