



毕业设计（论文）外文文献翻译

学 院： 计算机科学与技术学院

专业班级： 2019 级计算机科学与技术 2 班

姓 名： 张晓波

学 号： 1925103047

指导教师：
(签 名)

翻译文献（见附件）来源：

ChatGPT: The End of Online Exam Integrity

说明：

1. 文献翻译正文部分建议参考《计算机学报》格式排版，各专业可根据实际制作本专业的模板并发布。指导老师需审核英文翻译是否符合要求，并签名以示通过审核。没有老师签名的版本，视作不合格，不能装袋存档。
2. 文献原文作为附件放在翻译文档后。

ChatGPT：在线考试完整性的终结？

Teo Susnjak

数学与计算科学学院，梅西大学，新西兰奥克兰

2022 年 12 月 20 日

摘要 该研究评估了最近开发的人工智能（AI）代理 ChatGPT 执行高级认知任务并生成与人类生成的文本难以区分的文本的能力。这种能力引起了人们对在线考试中可能使用 ChatGPT 作为学术不端行为工具的担忧。研究发现，ChatGPT 能够展示批判性思维技能，并以最少的输入生成高度逼真的文本，这对在线考试的完整性构成了潜在威胁，特别是在高等教育环境中，在线考试越来越普遍。回到监考和口试可以成为解决方案的一部分，使用先进的监考技术和人工智能文本输出检测器也可能会有有效地解决这个问题，但它们都不太可能是万无一失的解决方案。我们需要进一步的研究来充分理解像 ChatGPT 这样的大型语言模型的含义，并制定策略来应对使用这些工具作弊的风险。为了维护在线考试对所有学生的公平性和有效性，教育工作者和机构必须意识到 ChatGPT 被用于作弊的可能性，并研究解决这一问题的措施。

关键词 ChatGPT；在线考试；大型语言模型；评估作弊；学术诚信；监考考试；监考工具；GPT-3

ChatGPT: The End of Online Exam Integrity?

Teo Susnjak

School of Mathematical and Computational Sciences, Massey University, Auckland, New Zealand

December 20, 2022

Abstract This study evaluated the ability of ChatGPT, a recently developed artificial intelligence (AI) agent, to perform high-level cognitive tasks and produce text that is indistinguishable from human-generated text. This capacity raises concerns about the potential use of ChatGPT as a tool for academic misconduct in online exams. The study found that ChatGPT is capable of exhibiting critical thinking skills and generating highly realistic text with minimal input, making it a potential threat to the integrity of online exams, particularly in tertiary education settings where such exams are becoming more prevalent. Returning to invigilated and oral exams could form part of the solution, while using advanced proctoring techniques and AI-text output detectors may be effective in addressing this issue, they are not likely to be foolproof solutions. Further research is needed to fully understand the implications of large language models like ChatGPT and to devise strategies for combating the risk of cheating using these tools. It is crucial for educators and institutions to be aware of the possibility of ChatGPT being used for cheating and to investigate measures to address it in order to maintain the fairness and validity of online exams for all students.

Keywords ChatGPT; online exams; large language models; assessment cheating; academic integrity; invigilated exams; proctoring tools; GPT-3

1 介绍

近年来，高等教育出现了向在线学习转变的情况，而 COVID-19 大流行加速了这一趋势^[7]。许多高等教育机构（HEIs）不得不通过过渡到在线课程和考试来迅速适应流感大流行带来的挑战^{[8][9][15]}。在线教育的这些情况不太可能在不久的将来被逆转^[12]。尽管遇到了挑战，

但远程学习的好处已经得到了高等教育机构和学生的认可^[7]。

随着教育越来越多地被转移到网上，人们对学术诚信的担忧也被放大了^{[7][14][25]}。特别是向在线考试的过渡，引起了人们对作弊和其他形式的学术不端行为的担忧^{[3][6][11][15][18]}。这在一定程度上是因为在线考试固有的匿名性和缺乏直接监督，以及学生在考试期间可以轻松访问和共享考试资源。

虽然在线考试中的学术诚信问题已经引起了人们的关注，但目前还缺乏明确的研究，也没有对在线评估中不诚实行为的程度进行量化的综合文献综述^[14]。有迹象表明，这项研究的流行率正在上升。在早期的研究中，Fask 等人^[13]，Corrigan-Gibbs 等人^[10]，Alessio 等人^[4]发现在线考试中作弊率很高，而在一般情况下，Arnold^[6]也提到在教育工作者中有一种观点，即学术不端行为呈上升趋势，而在线考试特别有利于作弊。最近，Noorbehbahani 等人^[18]报道称，他们关于在线考试作弊的研究已经进行了十多年，发现在线考试中的不诚实行为比传统的面对面考试更为普遍。

为了维护在线考试中的学术诚信，高等教育机构已经实施了修订后的评估建议^[25]，以及各种技术策略，如监考考试^[4]，抄袭检测软件，考试安全措施，同时还修订了机构学术诚信政策和教育活动，以阻止不当行为和维护荣誉规章制度^[10]。虽然这些策略单独或联合使用可能会有效降低学术不端行为的风险，但目前没有足够的证据来重新评估其在在线考试中维护学术诚信的整体有效性。同时，围绕在个人电脑上使用监考软件的道德问题^[5]以及他们最近在法律上的挑战^[2]也获得了发展的势头。

高等教育机构探索的另一项措施是转向使用更具挑战性的考试问题^[16]，这些问题需要更高层次的批判性思维。Whisenhunt 等人^{[25][1]}认为值得注意的是，这些类型的问题，包括论文和简短的答题，通常被教育工作者认为更适合测量批判性思维^[22]以及促进更深入的学习^{[24][26]}。它们背后的潜在意图是远离多项选择和简单的信息检索问题，因为当遇到涉及未经授权的 Web 访问的不当行为时，这些简单类型的问题被认为更容易作弊^[18]。

然而，如今在线考试的学术诚信已受到了新的威胁，即使是那些需要高阶推理的考试也是如此。最近的^[2]OpenAI 公开发布了 ChatGPT^[1]，世界已经见证了涉及自然语言处理和推理的人工智能能力的重大飞跃。这个公开的^[3]技术不仅能够进行复杂的对话，而且能够提供关于几乎所有主题的信息。对于需要高级分析、综合和应用信息的困难问题，它也能够产生令人信服和准确的答案，这将在本研究中得到证明。它甚至可以自己设计关键问题，这些问题正是不同学科的教育工作者用来评估学生能力的问题。假设高风险考试将继续被认为是具有价值的，并将继续在教育中使用，这一发展可能意味着在线考试的学术诚信的终结。因此，必须对这种人工智能代理的能力进行检查。

1.1 ChatGPT

ChatGPT 是一个大型语言模型。大型语言模型是一种使用深度学习（机器学习的一种形式）来处理和生成自然语言文本的人工智能。这些模型在大量文本数据上进行训练，使它们能够学习人类语言的细微差别和复杂性。在 ChatGPT 的案例中，它在各种文本数据（包括书籍、文章和在线对话）上进行了训练，以使其能够参与重要的对话，并就广泛的主题提供准确的信息，^[4]ChatGPT 的开发代表了自然语言处理和人工智能领域的重大进步，它建立在最初的 GPT 之上（生成预训练转换器^[23]），并为这一领域的进一步创新铺平道路。

¹ 作者制定了一套在线环境中进行多项选择考试的建议。附录 A 展示了 ChatGPT 回答此类问题的能力。

² ChatGPT 的测试版研究于 2022 年 11 月 30 日发布

³ 截至本文发表之日，该技术已向公众开放。

⁴ 用于训练 ChatGPT 的数据集的详细信息尚未公开发布。

这些大型语言模型的关键优势之一是它们能够理解给定提示的上下文并生成适当的响应。本文将重点演示这一功能。与早期的语言模型相比，这是一个显著的改进，早期的语言模型通常无法解释给定文本背后的含义和意图。另一个重要方面是它能够生成难以与人类书写区分的高质量文本，凭借其获取知识和回答困难学术问题的能力，它天生就能够回答那些通过网络搜索不容易回答的考试问题，并为其提供准确可靠的答案。

1.2 目标和贡献

本文的目的是对 ChatGPT 的能力进行初步分析，以展示其推理能力的复杂性及其跨多个学科回答非平凡的大学级别问题的能力，以及其清晰和连贯地呈现其回答的能力。

这项研究的贡献在于它的新颖性，以及对像 ChatGPT 这样的大型语言模型对学术诚信构成的前所未有的威胁提出警告的紧迫性。其目的是强调这样一个事实，即目前维护学术诚信的缓解策略、建议和程序可能还无法解决这一危险。

虽然这项技术可以说是自网络搜索引擎问世以来人工智能能力的最大进步之一，并将为生产力的提高提供巨大的机会，但它也给 HEI 部门带来了新的挑战，该部门已经处于混乱状态，不得不紧急研究这一发展。

2 背景

文献综述考察了对学术诚信问题的最新调查，更多地关注了在线评估的背景。

Butler-Henderson 和 Crawford^[8]进行了一项系统审查，强调了学习和教学向更积极的学习环境的转变，特别是在 COVID-19 大流行的背景下。这导致了在线考试形式的采用，作者认为这是由增加国际入学人数的愿望和高等教育的“大众化”所驱动的，而大流行病的影响是加速了这些趋势。作者指出了在线考试的局限性和挑战，包括作弊问题，以及技术的获取和缺乏标准化的方法。该研究最后呼吁进一步研究在线考试以及设计公平、有效和可靠的在线考试的重要性。

在一份综合报告中，Barber 等人^[7]讨论学术不端行为，包括剽窃和作弊，是高等教育机构和教育工作者在面对面和数字评估中关注的问题。技术在帮助机构检测剽窃行为方面发挥了重要作用，生物认证、作者身份分析和监考软件等技术的新发展也正被用于识别不当行为。然而，监考软件的使用也引发了对隐私的担忧，由于带宽的差异，国际学生在使用该软件时遇到了问题。在 COVID-19 大流行期间，向数字化教学和学习的转变促使人们对评估方法进行审查，包括使用开卷在线考试和更真实的综合评估。

Coghlan 等人^[9]报道称，使用人工智能和机器学习的在线考试监考技术已经获得了关注。然而，该研究指出，这些技术面临着争议和伦理问题，包括学生隐私、潜在偏见以及软件的有效性和可靠性等问题。一些大学为这些技术的使用进行了辩护，而另一些大学则放弃或拒绝使用这些技术。

Noorbehbahani 等人^[18]最近对网上考试作弊进行了系统的文献综述，其涵盖了十多年的研究发现，在线考试作弊确实是一个值得关注的问题。该研究称，在线考试中的作弊现象比传统的面对面考试更为普遍。作者指出，可以使用广泛的技术和工具来促进作弊，例如远程桌面和屏幕共享，在互联网上搜索解决方案，以及使用社交网络。除了在线监考，作者还确定了一种潜在有效的预防策略（如防作弊问题）和检测方法（如剽窃检测软件和机器学习算法）的组合。

Henderson 等人^[15]还发现，在线考试中作弊的盛行是一个重要问题，而即使在校园内的纸质监考考试中，这仍然是一个问题。作者的发现指向了之前的研究，该研究表明，尽管采取了安全措施，但作弊现象依然存在，关于监考和网络安全对作弊的影响，现有的证据相互矛盾。他们的结论是，虽然基于技术的安全措施可以影响学生的体验和对诚信的态度，但它们并不一定会减少作弊。

关于监考软件，Alin 等人^[5]还提出了围绕在个人计算机上使用该技术的道德问题，以及对构成可疑行为的解释。作者强调，即使允许在线考试中使用监考系统，考试仍然容易作弊。作者认为，目前人们对虚拟监考考试中如何作弊以及如何最好地减少作弊缺乏了解。同时，可汗等人^[16]还强调，监考软件要求学生在线考试期间保持摄像头打开，这被学生认为是压力和对隐私的侵犯，同时他们还认为，无论采取何种措施，作弊都将继续。该研究还建议使用一些策略，如用简答题代替多项选择题，并采用更严格的时间限制。而 KOH 和 Daniel^[17]还注意到，在教学过渡到在线模式时，教育工作者使用的关键策略之一是将多项选择题转换为书面批判性思维问题，学生们报告称，这一举措导致他们认为考试变得更难。

3 方法论

这里描述了检查 ChatGPT 的批判性和高阶思维能力的方法。下面列出了三个步骤，并在本节中进行了更详细的描述。

1. 首先，ChatGPT 被要求自己生成涉及一些场景的批判性思维难题的例子，这些问题针对来自不同学科的本科生。
2. 其次，要求 ChatGPT 回答生成的问题。
3. 最后，ChatGPT 被要求对问题的答案进行批判性评估。



图 1 ChatGPT 的可公开访问的在线界面，底部显示了文本输入提示

设置：通过 OpenAI 创建了一个帐户^[19]，在 Beta 研究版本下授予对 ChatGPT 的访问权限，以进行实验性使用。图 1 显示了访问 ChatGPT 的在线门户。

所选学科：从科学、教育研究、人文学科和商业中选择了广泛的学科领域用于演示目的。具

体来说，ChatGPT 被提示生成关于机器学习、市场营销、历史和教育等特定学科的特定主题问题和回答。

问题生成：为了测试 ChatGPT 跨多个学科生成批判性思维问题的能力，该模型首先提供了一个提示，指出每个研究领域都有一个批判性思维问题的例子，这些问题具有挑战性，但适合本科生。这是通过为每个学科提出以下提示来实现的：

“请生成一个针对本科生的机器学习领域的困难关键问题的示例，其中涉及一个假设场景。”

答案生成：随后提示 ChatGPT 回答其为每个学科设计的问题，并对回答的结构和形式进行具体说明，如下所示

“请用 500 字分几段回答下面的问题，并举例和论证。”

对响应的批判性评估：为了测试 ChatGPT 对其生成的响应进行批判性评估和分析的能力，并提供改进建议，提示模型考虑提出的原始问题和给出的答案，并提供以下附加说明：

“能否请您批判性地评价以下对上述问题的回答，列出其优点和缺点以及改进建议。”

3.1 评价答复

通用智力标准是适用于评估任何问题、议题或情况的推理质量的标准^[21]，其所表达的批判性思维水平应该考虑所提出的论点的目的、信息和含义，并且需要使用 Paul 所概述的通用智力标准来评估这些论点的质量^[20]来增加说服力和独创性。为此，以下代表批判性思维维度的标准将应用于回答，以评估 ChatGPT 对上述提示提供的回答的质量。

相关性：所表达的观点是否与当前的主题或问题相关？它是否解决了问题的复杂性？

清晰度：文本是否易于理解？它是否结构合理、逻辑清晰？它是否为目标受众提供了适当的语言和词汇？

准确性：所表达的思想是真实的、准确的？是否可以通过证据或其他方式进行验证？

精确：表达的想法是否足够具体和详细？它是精确而不含糊的吗？

深度：所表达的想法是否超越了表面层次，并考虑了问题背后的复杂性和细微差别？文本是否对主题进行了全面深入的分析？它是否考虑了多个角度，并提出了一个平衡的观点？

广度：所表达的想法是否考虑了关于该问题的所有相关观点和看法？

逻辑：所表达的思想是否遵循逻辑一致的推理？结论是否得到证据的支持？

说服力：文本是否有效地说服读者接受其论点或结论？提出的证据是否有力和令人信服？

原创性：文本是否提供了新的见解或想法，还是仅仅重复了已经广为人知的信息？

4 结果

ChatGPT 对方法中概述的提示的响应显示在四个表中，每个表代表一个单独的学科，即

表 1 中的教育，表 2 中的机器学习，表 3 中的历史学科和表 4 中的市场营销。

4.1 响应分析

在对再生响应进行高阶临界分析时，有几个关键因素需要考虑。

清晰度：在对提示的所有响应中，ChatGPT 都表现出了很强的清晰度。它回答中使用的语言都易于理解，并且遵循了自然语言的结构和惯例。它的回答是组织良好并且连贯的，而且在较长的文本表达中我们可以感受到一种有意的思想流动。同时，它为创造以下这些问题提供的理由以及最后重要的自我评价中也都表现出了它的明确性。而在词汇表达上，尤其是在必要的技术词汇以及语法上，它的表达结果也都是可以被目标受众所接受的。

表 1：由 ChatGPT 生成的批判性思维问题的示例及其对教育学科的基本推理。对 ChatGPT 提出的问题的回答，以及 ChatGPT 自己对其回答的批判性评价和改进建议。

有道理的问题	回答
<p>“在假设的情况下，如果一个学区的学生缺勤率很高，那么实施个性化学习计划会如何影响出勤率和学生成绩，以及在设计和实施此类计划时应考虑哪些因素？”</p> <p>理由：上述问题是合适的，因为它提出了一个与教育领域相关的假设情景，并挑战本科生批判性地思考如何解决现实世界的问题。该问题要求学生考虑个性化学习计划对出勤率和学生成绩的潜在影响，以及在设计和实施此类计划时应考虑的因素。这种类型的问题鼓励学生将他们对教育理论和实践的知识和理解应用到假设的情况中，并考虑不同方法的潜在后果。总的来说，这个问题是合适的，因为它挑战学生对教育领域的相关问题进行批判性和创造性的思考。</p>	<p>在假设的情况下，一个学区正在经历高学生缺勤率，实施个性化学习计划可能会对出勤率和学生成绩产生积极影响。个性化学习计划旨在根据每个学生的个人需求和兴趣定制教学，这可以提高学习的积极性和参与度。当学生觉得他们的学习是有意义的，并且与他们自己的生活相关时，他们更有可能定期上学，并在学习中付出努力。</p> <p>有证据支持个性化学习计划可以提高出勤率和学生成绩的观点。例如，美国教育部的一项研究发现，与参加传统课程的同龄人相比，参加个性化学习课程的学生出勤率更高，学习成绩也更好。同样，Clayton Christensen 研究所对个性化学习研究的回顾发现，个性化学习计划对学生的参与度、动机和成就有积极影响。</p> <p>然而，在学生缺勤率较高的学区设计和实施个性化学习计划时，还需要考虑一些因素。一个考虑因素是需要足够的资源和支持。个性化学习计划通常需要教师给予更多的个性化关注，并且可能需要额外的技术和材料。重要的是要确保学校拥有必要的资源和支持，以有效地实施个性化学习计划。另一个考虑因素是教师需要适当的培训和专业发展。个性化学习方法可能不同于传统方法，教师可能需要支持和指导才能在课堂上有效地实施这些方法。</p> <p>总体而言，在学生缺勤率高的学区实施个性化学习计划可能会对出勤率和学生成绩产生积极影响。然而，重要的是要仔细考</p>

	虑有效实施这一计划所需的资源和支持，以及教师的培训和专业发展需求。通过考虑这些因素，学区可以增加通过个性化学习提高出勤率和学生成绩的成功机会。
<p>对答案的批判性评价：</p> <p>优势：回答提供了证据来支持个性化学习计划可以提高出勤率和学生成绩的观点。它承认，为了有效实施个性化学习方案，需要为教师提供充足的资源和支持以及适当的培训和专业发展。</p> <p>缺点：回答没有提供关于所提到的研究或调查的具体例子或细节。了解更多关于这些研究的背景和方法的信息将有助于更好地理解他们的研究结果以及如何将其应用于假设情景。该回答没有解决在学生缺勤率高的学区，实施个性化学习计划的任何潜在挑战或缺点。考虑可能出现的潜在障碍或困难以及如何解决这些障碍或困难将是有益的。</p> <p>改进建议：提供关于所提及的研究或调查的具体例子和细节，包括关于所使用的背景和方法的信息，这将有助于为调查结果提供更多的背景和可信度。考虑在学生缺勤率高的学区实施个性化学习计划的潜在挑战或缺点，并提出解决策略，这将为个性化学习的潜在影响提供更加平衡和细致入微的视角。在此假设情景中，考虑可能影响出勤率和学生成绩的其他因素，如社会经济因素、学校文化或社区资源，这将对这一问题进行更全面的分析，并提出更全面的解决办法。</p>	

表 2：由 ChatGPT 生成的批判性思维问题的示例及其对机器学习学科的基本推理。对 ChatGPT 提出的问题的回答，以及 ChatGPT 自己对其回答的批判性评价和改进建议。

有道理的问题	回答
<p>“想象一下，你正在进行一个机器学习项目，以预测客户从订阅服务流失的可能性。你的模型在训练数据上表现良好，但当你在测试数据上测试它时，准确性明显下降。出现此问题的可能原因是什么？您将如何解决此问题？”</p> <p>理由：上述问题是合适的，因为它提出了一个在现实世界的机器学习项目中经常遇到的假设场景。它要求学生批判性地思考可能影响机器学习模型性能的各种因素，并考虑如何排除故障和解决任何问题。此外，这个问题还涉及到在测试数据上评估模型的重要性，这是机器学习过程中的关键一步。总体而言，该问题要求学生将他们的机器学习原理和技术知识应用到现实场景中，这有助于加深他们对该领域的理解。</p>	<p>训练数据和测试数据之间的性能差异的一个可能原因是过拟合。当模型过于复杂，并且能够非常好地拟合训练数据，但不能推广到新数据时，就会发生过拟合。这可能是由于模型中的参数太多，或者训练数据量有限造成的。在任何一种情况下，模型对于训练数据变得过于专门化，并且不能准确地预测新数据的结果。</p> <p>要解决此问题，一种方法是尝试通过减少参数数量或增加训练数据量来简化模型。这可以帮助模型更准确地捕捉数据中的潜在模式，并提高其泛化性能。另一种选择是使用诸如正则化之类的技术，这有助于通过引入在模型中使用过多参数的惩罚来防止过拟合。如果过度拟合不是差异的原因，则可能有其他因素在起作用，一种可能性是训练数据和测试数据来自不同的分布。例如，训练数据可以从特定时间段或地理区域收集，而测试数据可以来自不同的时间段或区域。在这种情况下，模型可能在训练数据上表现良好，因为它已经学习了特定于该数据的模式，但这些模式可能无法推广到测试数</p>

	<p>据上。</p> <p>为了解决这个问题，一种解决方案可以是确保训练数据将代表应用模型的数据的总体分布。这可以通过诸如分层采样之类的采样技术来实现，其确保训练数据包含总体数据内的不同子组的平衡表示。另一种选择是使用领域适应技术，这有助于模型学会适应数据分布的变化。</p> <p>总的来说，训练数据和测试数据之间的性能差异有许多潜在的原因，适当的故障排除方法将取决于项目的具体情况。通过仔细评估模型和数据，并应用适当的技术来解决问题，可以提高模型的泛化性能，并提高其对新数据的准确性。</p>
<p>对答案的批判性评价：</p> <p>优势：回答正确地将过度拟合识别为训练数据和测试数据之间性能差异的可能原因。回答提供了解决该问题的几个建议，包括简化模型、使用正则化技术以及确保训练数据代表整体数据分布。</p> <p>缺点：除了过度拟合和数据分布之外，回答没有提及差异的其他可能原因，例如训练数据中的偏差或模型中缺乏相关特征。回答没有提供可用于简化模型或确保训练数据代表总体数据分布的技术的具体示例。</p> <p>改进建议：提及差异的其他潜在原因，例如训练数据中的偏差或模型中缺乏相关特征。提供可用于简化模型或确保训练数据代表整体数据分布的技术的具体示例，例如特征选择或采样技术。解释建议的每种技术的优点和缺点，并就如何决定哪种技术最适合给定的情况提供指导。</p>	

表 3：由 ChatGPT 生成的批判性思维问题的示例及其对历史学科的基本推理。对 ChatGPT 提出的问题的回答，以及 ChatGPT 自己对其回答的批判性评价和改进建议。

有道理的问题	回答
<p>“想象一下，一个新发现——有一件事完全改变了我们对一个重大历史事件的理解。这一发现将如何影响我们对过去的理解，以及需要采取哪些步骤才能将其纳入历史叙事？”</p> <p>理由： 这个问题需要深刻理解历史叙事在塑造我们对过去的理解中所起的作用，以及批判性地思考重大历史发现的潜在影响和意义的能力。它还需要认识到不同的观点和解释会如何影响我们对历史的理解，并有能力在假设情景的背景下提出应对这些挑战的潜在解决方案。</p>	<p>一项新发现彻底改变了我们对重大历史事件会深刻影响我们对过去的认识。这一发现可能会挑战长期以来对这一事件的看法和假设，并要求历史学家重新评估和修正他们对过去的解释。</p> <p>这类发现的一个例子可能是关于世界历史上一个关键事件的新证据的发现，比如 John F.Kennedy 总统遇刺。几十年来，这次暗杀一直被认为是一名单独的枪手所为，但想象一下，如果发现了新的证据，表明有一个更广泛的阴谋在起作用。这一新信息将从根本上改变我们对这一事件及其历史地位的理解。</p> <p>要将这一发现纳入历史叙述，需要采取一系列步骤。首先，历史学家需要仔细审查</p>

	<p>和评估新证据，以确定其真实性和相关性。这一进程可能涉及进行更多的研究和分析，以及寻求该领域其他专家的投入。如果发现证据可信，则有必要将其纳入现有的历史叙述中。</p> <p>这一过程将涉及修改对事件的现有解释，并发展新的理论来解释新的信息。这可能涉及根据新发现重新审视其他历史事件和来源，以及考虑该发现如何改变我们对更广泛历史背景的理解。此外，重要的是将新发现传达给公众和其他利益攸关方，如教育工作者和决策者，以确保修订后的历史叙述得到广泛理解和接受。</p> <p>总的来说，一项新发现彻底改变了我们对重大历史事件的理解，将我们对过去的理解产生重大影响。通过仔细评估并将这些新信息纳入历史叙述，我们可以确保我们对过去的理解是准确和完整的。</p>
<p>对答案的批判性评价：</p> <p>对这个问题的回答写得很好，详细讨论了一个新发现的潜在影响，它完全改变了我们对一个重大历史事件的理解。回答清楚地解释了这一发现将如何挑战现有的信念和对过去的解释，并对将这一发现纳入历史叙述所需采取的步骤进行了深思熟虑的分析。</p> <p>回答的一个优势是它使用了一个具体的例子来说明新发现的潜在影响。通过讨论有关 John F.Kennedy 总统遇刺的新证据的假设发现，回答提供了一个具体的例证，说明新发现可能会如何挑战现有的历史叙述，并要求历史学家修正他们对过去的解释。</p> <p>回答的另一个优点是它的彻底性。回答涵盖了将新发现纳入历史叙述所需采取的所有关键步骤，从仔细评估证据到修改现有解释，以及将新信息传达给更广泛的公众。</p> <p>回答的一个潜在缺点是，它没有讨论在将新发现纳入历史叙述的过程中可能出现的潜在挑战和困难。例如，回答没有提到其他历史学家或公众可能对现有历史叙述的抵制或怀疑。此外，该回答并未涉及新发现对更广泛问题的潜在影响，如历史教学或公共政策。</p> <p>为了改进回答，一项建议是提供更多的细节和分析，说明在将新发现纳入历史叙述的过程中可能出现的潜在挑战和困难。这可能包括讨论来自其他历史学家或公众的潜在阻力或怀疑，以及新发现对历史教学或公共政策等更广泛问题的潜在影响。此外，通过提供更具体的例子，说明新发现如何挑战现有的历史叙述，并要求历史学家和对过去的解释，可以加强回答。</p>	

表 4：由 ChatGPT 生成的批判性思维问题的示例及其对市场营销学科的基本推理。对 ChatGPT 提出的问题的回答，以及 ChatGPT 自己对其回答的批判性评价和改进建议。

有道理的问题	回答
“想象你是一名营销人员。一家销售有机护肤品的公司的经理。一家竞争公司刚刚发布了一系列新的合成护肤品，价格便宜得多，	在一家销售有机护肤品的公司担任营销经理，面对来自更便宜和更广泛可用的合成产品的竞争，这是一个令人生畏的挑战。

而且在商店里更容易买到。你的目标受众主要是注重健康的千禧一代女性。你将如何制定营销策略来区分你的有机产品，并说服消费者选择你的品牌而不是更便宜的合成产品？”

理由：这是一个合适的问题，因为它提出了营销经理在现实世界中可能面临的实际场景，并要求学生批判性地思考如何解决这个问题。它还要求学生考虑目标受众和他们的动机，以及有机和合成产品的优点和缺点。该问题还允许多种可能的解决方案，鼓励学生考虑各种策略，并考虑每个选项的潜在结果。总的来说，这个问题要求学生将他们的知识和技能应用于市场营销领域的现实和复杂的问题。

然而，通过精心策划和有针对性的营销策略，有可能区分我们的有机产品，并说服消费者选择我们的品牌而不是合成产品。

一种可能的方法是着重强调使用有机产品的好处。许多消费者，特别是千禧一代的女性，越来越关注她们的购买决定对环境和个人健康的影响。通过强调我们产品的可持续性和天然成分，我们可以迎合这一群体的价值观，并在合成竞争中脱颖而出。例如，我们可以利用社交媒体活动和有影响力的合作伙伴关系来展示使用有机产品的好处以及合成成分对环境和个人健康的负面影响。

此外，我们还可以利用我们的独特卖点（USP）进一步在竞争中脱颖而出。例如，如果我们的产品由稀有或难以找到的有机成分制成，或者如果我们提供比合成品牌更广泛的产品，我们可以在营销工作中强调这些要点。通过专注于使我们的品牌脱颖而出的因素，我们可以让消费者相信我们的产品值得付出额外的成本。

另一种策略可能是在更大的有健康意识的千禧一代女性群体中瞄准利基市场。例如，我们可以关注孕妇或皮肤敏感的女性，她们可能特别关注护肤品中的成分。通过针对目标受众的特定子集定制我们的营销工作，我们可以有效地将我们产品的优势传达给更有针对性的消费者群体。

最后，正面解决消费者对产品成本的担忧可能也是有益的。一种方法是通过使用忠诚度计划或特别优惠，如首次购买者的折扣或捆绑交易。通过为我们的客户提供价值，我们可以帮助抵消我们产品的高成本，并使其对价格敏感的消费者更具吸引力。

总体而言，有许多潜在的方法来区分我们的有机护肤品，并说服消费者选择我们的品牌，而不是更便宜的合成产品。通过强调使用有机产品的好处，利用我们的 USP，瞄准利基市场，并解决成本问题，我们可以有效地传达我们产品的价值，并说服消费者进行转换。

对答案的批判性评价：

上面对这个问题的回答提出了几个潜在的策略来区分有机护肤品，并说服消费者选择它们而不是更便宜的合成产品。回答的一个优势是，它考虑了目标受众的动机和价值

观，特别是具有健康意识的千禧一代女性，并建议针对这些价值观开展营销工作。它还强调了利用独特卖点和解决成本问题作为区分有机产品的方法的重要性。

然而，应对措施中有一些不足之处，可以加以改进。一个缺点是，回答没有考虑到实施所建议的战略的潜在挑战或缺点。例如，瞄准利基市场可能在短期内更有效，但也可能限制品牌的长期增长潜力。此外，提供折扣或忠诚度计划可能有助于吸引对价格敏感的消费者，但也可能侵蚀产品的感知价值。

另一个缺点是，回答没有考虑来自合成品牌的潜在竞争。例如，合成品牌也可以针对有健康意识的消费者或提供他们自己的忠诚度计划。通过考虑竞争格局，营销经理可以更好地预测竞争对手的行动，并相应地调整他们的战略。

总体而言，回答提出了几种潜在的策略，以区分有机护肤品，并说服消费者选择它们而不是更便宜的合成产品。为了改进回答，考虑每种战略的潜在挑战和缺点以及竞争格局将是有益的。

准确性：为了充分考虑对问题的答复的准确性，需要征求四个学科的学科专家的评价。作者可以证明回答和问题的准确性，以及随后关于机器学习的回答的评价很好地描述了过度拟合的概念，并准确地提供了可用于解决该问题的技术示例。

从市场营销、教育（专注于美国背景）和历史领域的学科专家来评估这些回答的准确性超出了本研究的范围。回到机器学习问题，它提出的问题与作者在数据科学课程中使用的问题相同。在假设的情况下，提出这个问题的理由也是正确的，对所产生的反应的评判也是敏感的，并且如果整合到实际的考试反应中，这将会得到满分，而在作者的经验中，实际的学生很少能得到满分。

精确：对问题的回答既具体又详细。在机器学习回答的上下文中，提供了可用于解决问题的技术的具体示例。回答还清楚地区分了差异的不同潜在原因。在教育方面，美国教育部和 Clayton Christensen 研究所的研究被引入了回答。在历史方面，讨论了刺杀 John F. Kennedy 总统的例子，而营销的回答则确定了具体的目标群体。在所有关键评估中，它都提供了具体和详细的要点，也证明了回答的精确性。

相关性：通常，对于每个学科的每组请求，为生成初始考试问题的请求提供的响应，接着是答案，随后是对答案的批判性评估，都与主题明显相关。所有的回答也都与主题相关，也与每个学科的主题和请求的意图相关。

深度：ChatGPT 在所提出的问题的复杂性以及为支持这些问题而提供的理由方面表现出了值得注意的深度。答案限制在 500 字以内，在此范围内，为所有学科提供了经过深思熟虑的策略和示例。同时，对回答的评论是实质性的，有优点和缺点，以及根据要求提出的改进建议，证明了人工智能代理对每个主题的全面和深入的分析理解。

广度：同样，考虑到需要给出回答的限制条件，它所提供的答案在每种情况下都提供了两种情况的解释。而每一种情况下的改进建议都提供了进一步的例子，说明代理的广度能力。

逻辑：所有的回答都遵循逻辑一致的推理，提供具体的例子和解释。在基本原理中特别强调逻辑的连贯性。以证明每个问题的生成，以及在对相同问题的回答中，对不同的场景进行了解释和组织。

说服力：虽然学科专家可能会对人工智能代理提供的答案提出额外的批评意见，但回答的一个显著特征是他们表达的信心。论点和证据以清晰和合乎逻辑的方式提出，并提到了解决潜在的反驳或反对意见的努力。最终，回答的说服力的质量将根据读者的观点而有所不同，这是不言而喻的，无论对错，都是毫无保留地表达出来的。

原创性：人工智能代理已经在包含大量积累知识的人类生成文本的大型数据集上进行

了训练，并通过设计，根据该数据集中发现的模式和惯例生成响应。因此，如果回答是正确的，很难想象这些回答是原创的。然而，由于准确性是大型语言模型的已知弱点，因此可以想象，在模型的训练过程中，来自数据集语料库的主题和事件之间的一些连接是不正确的（称为“幻觉”），在这种情况下，如果它们是不正确的，则一些响应可能是原始的。在这种情况下，所提供的回答很可能提供了基于语言模型及其应用领域的已有知识和实践的信息的见解，虽然这些回答在提出全新想法的意义上可能不一定是原创或新颖的，但它们提供了看似有用的信息和见解，这足以回答本科考试。

5 讨论

从本文中进行的实验证据中可以清楚地看到，人工智能技术已经达到了非凡的水平，现在它已经能够进行批判性思维，而不仅仅是信息检索。所产生的回答可以被评估为阐述清楚，所用例子准确，与请求相关，同时考虑到所施加的限制，具有足够的深度和广度，同时在较长的文本中具有逻辑上的连贯性。人工智能技术的这些进步不仅仅是渐进式的改进，而且可以被视为过去 25 年来最重大的技术飞跃。

ChatGPT 最令人印象深刻的能力之一是它的批判性推理能力，以及用完美的散文表达思想和观点的能力。这项技术在这些领域表现出了非凡的能力，似乎与人类的能力相当。Paul^[20]指出一个人如果是一个批判性思考者，在某种程度上，他们会定期学习和批评自己的想法以提高它。人工智能智能体能够对自己的反应进行评论，并对其优点和缺点进行明确的讨论，并提出改进建议，这表明智能体具有这种批判能力。人工智能代理已经超越了当前网络搜索引擎所提供的单纯的记忆和信息检索能力，相反，现在能够概念化和展示更高层次的思维。

随着时间的推移，这些进步对社会的全面影响将变得明显，但对教育部门的影响是深远的，特别是与在线考试的完整性直接相关。学生可能会使用 ChatGPT 在考试中作弊，因为该技术能够产生与人类难以区分的反应。本研究的实验结果支持这一观点，这对未来在线考试的可靠性和有效性以及作弊的可能性提出了严重质疑。

总体而言，ChatGPT 在批判性思考和推理方面的能力，以及用高质量的散文表达思想和观点的能力，既带来了机遇，也带来了挑战。虽然它具有促进思想交流、提高生产力和促进学习的潜力，但它也给教育部门带来了挑战。这项技术将需要仔细考虑和监督，以及就什么是合乎道德和负责任地使用这项技术达成共识。

5.1 缓解战略建议

最后，当 ChatGPT 被问到“你建议如何防止学生使用 ChatGPT 在大学在线考试中作弊？”时，给出了以下回答：

“为了防止学生使用 ChatGPT 或任何其他人工智能工具在在线考试中作弊，大学可以采取一些措施。

首先，大学可以采取一些措施防止学生在考试期间使用 ChatGPT 或任何其他类似工具。例如，他们可以使用监考软件来监控学生的屏幕，并确保他们在考试期间没有使用任何未经授权的工具或资源。他们还可以使用安全浏览器或其他技术来防止学生在考试期间访问外部网站或应用程序。

其次，大学也可以采取措施来检测和阻止那些试图使用 ChatGPT 或任何其他人工智能工具在考试中作弊的学生。例如，他们可以使用作弊检测软件来识别和标记任何与 ChatGPT 或任何其他人工智能工具生成的答案可疑相似的考试答案。他们还可以使用机器学习算法来分析考试答案，并识别可能表明使用未经授权的工具或资源的任何模式或异常。

第三，大学也可以教育学生使用 ChatGPT 或任何其他人工智能工具在考试中作弊的道德含义。他们可以向学生提供有关大学学术诚信政策的信息，并提醒他们在考试期间使用未经授权的工具或资源的后果。这可以帮助防止学生试图使用 ChatGPT 或任何类似的工具在考试中作弊，并可以鼓励他们诚实和道德地对待考试。

除了上面提到的步骤，大学还可以考虑使用人工智能来帮助防止学生使用 ChatGPT 或任何其他人工智能工具在在线考试中作弊。例如，他们可以使用机器学习算法来分析考试答案，并识别可能表明使用未经授权的工具或资源的任何模式或异常。他们还可以使用自然语言处理（NLP）算法来检测任何与 ChatGPT 或任何其他人工智能工具生成的答案可疑相似的答案。通过这种方式使用人工智能，大学可以提高检测和阻止试图使用 ChatGPT 或任何其他人工智能工具在考试中作弊的学生的能力。这可以帮助确保考试过程的公平性、有效性和完整性，并可以防止学生获得相对于同龄人的不公平优势。

此外，大学也可以考虑使用人工智能来帮助开发更安全、更强大的在线考试平台。例如，他们可以使用人工智能来提高考试平台的安全性，并防止学生在考试期间访问未经授权的工具或资源。他们还可以使用人工智能来改善考试平台的用户体验，并使学生更容易在任何地点参加考试。

总体而言，大学可以通过以上几种方式使用人工智能来帮助防止学生使用 ChatGPT 或任何其他人工智能工具在在线考试中作弊。通过使用人工智能来提高考试平台的安全性和可靠性，并检测和阻止作弊，大学可以确保其考试过程的公平性、有效性和完整性。”

我们也许在上面发现了 ChatGPT 响应的弱点的一些小迹象，公平地说，这可能更多的是由于设计解决方案的困难所说的问题比它的能力更重要。首先，使用作弊检测软件的建议不太可能有效，因为 ChatGPT 能够在相同的提示下生成唯一的答案，尤其是在对输入进行

小扰动的情況下。這意味着 ChatGPT 生成的答案可能与其他任何答案都不相似，因此作弊检测软件很难将其识别为可疑答案。

其次，使用机器学习算法和自然语言处理（NLP）算法来检测可疑答案的建议也符合一些高校目前正在使用的方法^[14]，但如果需要开发这些解决方案，可能不现实。对于许多大学来说，这种解决方案可能过于昂贵，负担不起。这些技术需要大量的资源和专业知识来实施和维护，对大多数机构来说可能并不可行。有一些迹象表明，现有的 GPT-文本输出检测器具有识别 AI 生成的文本的能力^[5]，这是由于所有文本中的底层签名的存在。然而，这些工具需要进一步研究。

关于使用人工智能来提高在线考试平台的安全性和可靠性，以及检测和阻止作弊的建议可能无效。虽然人工智能在这些领域是一个有用的工具，但它不是万能药，可能无法完全解决使用 ChatGPT 或任何其他人工智能工具作弊的问题。不幸的是，询问 ChatGPT 是否生成了特定的文本片段以发现作弊行为似乎并不有效，因为使用此策略的初步尝试表明，它不会保留其在以前会话中生成的文本的记录。

第三，教育学生使用 ChatGPT 或任何其他人工智能工具在考试中作弊的道德含义的建议不太可能有效地防止作弊。虽然教育是促进学术诚信的重要组成部分，但它本身可能不足以阻止那些决心作弊的学生，而且这些举措已被证明收效甚微^[10]。ChatGPT 的一个众所周知的限制是它的单模态输入能力，这意味着 ChatGPT 只能接受人类文本作为输入。因此，在没有有效监考软件的在线考试中，考试就需要在表达问题时不要只是用文本。因此，以下策略可以考虑：

- 将多模态通道用于考试问题：将图像嵌入到考试问题中可以使學生更难作弊，并使 ChatGPT 更难生成准确的响应，因为该技术仅依赖于文本输入。
- 尝试使用预先录制的视频录制问题，将口头问题与图像相结合：这可以为试图作弊的學生增加额外的难度，并使 ChatGPT 产生准确的回答更具挑战性。
- GPT 输出检测：在各种门户网站上根据 GPT 语言检测器模型在线检查响应^[6]
- 回到口试：要求學生在网上或校园内实时口头展示他们的知识。

大型语言模型发展成为更通用的语言模型只是时间问题。迟早会出现具有整合多个渠道能力的人工智能代理^[7]，包括图像、视频和音频输入。就目前而言，利用技术的局限性是保持领先的唯一途径，同时还要仔细评估不同策略的有效性，并根据需要不断调整和完善它们。

5.2 局限性

这是对 ChatGPT 用于回答在线环境中的批判性思维问题的能力初步调查。因此，可以进行进一步的改进，例如使用来自教育、历史和市场营销的独立学科专家来评估回答，事实上，在未来的研究中可以使用学科专家和来自各种课程的以前的考试问题。即使这样，这项研究也证明了 ChatGPT 能够产生有效的问题的事实。

6 结论

本研究调查了最近发布的人工智能代理 ChatGPT 执行高阶思维任务的能力，以及生成

⁵ <https://huggingface.co/openai-detector/>

⁶ <https://huggingface.co/openai-detector/>

⁷ GPT-4 正在开发中，据报道正在研究多模式输入能力。

与人类难以区分的文本的能力，并将其用作在线考试中学术不诚实的工具。同时人工智能代理还被提示生成问题，提供理由，然后是答案和自我评价。

研究发现，像 ChatGPT 这种技术的出现对在线考试的完整性构成了重大威胁，特别是在在线考试变得越来越普遍的高等教育背景下。这些模型展示了高度的批判性思维，能够用很少的输入生成高度逼真的文本，使学生在考试中作弊成为可能。

恢复监考和口试以及使用先进的监考工具可能会有效地应对这一威胁，尽管它们并不是完美的解决方案。我们需要研究新的人工智能和机器学习工具，这些工具能够检测类 ChatGPT 模型的文本输出。同时还需要进一步的研究来充分理解这些大型语言模型的含义，并制定策略来解决使用这些工具作弊的可能性。为了维护在线考试的完整性，并确保对所有学生进行公平有效的评估，教育工作者和机构必须意识到这一工具可能会助长作弊，并且必须及时采取措施打击作弊。

参考文献

- [1] OpenAI. URL <https://openai.com/>.
- [2] Case 1:21-cv-00500-kdb document 37 aaron m. ogletree, plaintiff v. cleveland state university, defendant. URL <https://law.justia.com/cases/federal/district-courts/ohio/ohndce/1:2021cv00500/275187/37/>. Accessed on December 16, 2022.
- [3] K. Ahsan, S. Akbar, and B. Kam. Contract cheating in higher education: a systematic literature review and future research agenda. *Assessment & Evaluation in Higher Education*, pages 1–17, 2021.
- [4] H. M. Alessio, N. Malay, K. Maurer, A. J. Bailer, and B. Rubin. Examining the effect of proctoring on online test scores. *Online Learning*, 21(1):146–161, 2017.
- [5] P. Alin, A. Arendt, and S. Gurell. Addressing cheating in virtual proctored examinations: toward a framework of relevant mitigation strategies. *Assessment & Evaluation in Higher Education*, pages 1–14, 2022.
- [6] I. J. Arnold. Cheating at online formative tests: Does it pay off? *The Internet and Higher Education*, 29:98–106, 2016.
- [7] M. Barber, L. Bird, J. Fleming, E. Titterington-Giles, E. Edwards, and C. Leyland. Gravity assist: Propelling higher education towards a brighter future - office for students, 2021. URL <https://www.officeforstudents.org.uk/publications/ravity-assist-propelling-higher-education-towards-a-brighter-future/>.
- [8] K. Butler-Henderson and J. Crawford. A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity. *Computers & Education*, 159:104024, 2020.
- [9] S. Coghlan, T. Miller, and J. Paterson. Good proctor or “big brother”? ethics of online exam supervision technologies. *Philosophy & Technology*, 34(4):1581–1606, 2021.
- [10] H. Corrigan-Gibbs, N. Gupta, C. Northcutt, E. Cutrell, and W. Thies. Deterring cheating in online environments. *ACM Trans. Comput.-Hum. Interact.*, 22(6), sep 2015. ISSN 1073-0516. doi: 10.1145/2810239. URL <https://doi.org/10.1145/2810239>.
- [11] C. Crook and E. Nixon. How internet essay mill websites portray the student experience of higher education. *The Internet and Higher Education*, 48:100775, 2021.
- [12] D. Dumulescu and A. I. Muțiu. Academic leadership in the time of covid19—experiences and perspectives. *Frontiers in Psychology*, 12:648344, 2021.
- [13] A. Fask, F. Englander, and Z. Wang. Do online exams facilitate cheating? an experiment designed to separate possible cheating from the effect of the online test taking environment. *Journal of Academic Ethics*, 12(2):101–112, 2014.
- [14] M. Garg and A. Goel. A systematic literature review on online assessment security: Current challenges and integrity strategies. *Computers & Security*, 113:102544, 2022.
- [15] M. Henderson, J. Chung, R. Awdry, M. Mundy, M. Bryant, C. Ashford, and K. Ryan. Factors associated with online examination cheating. *Assessment & Evaluation in Higher Education*, pages 1–15, 2022.

- [16] S. Khan, M. E. K. Kambris, and H. Alfalahi. Perspectives of university students and faculty on remote education experiences during covid-19-a qualitative study. *Education and information technologies*, 27(3):4141–4169, 2022.
- [17] J. H. L. Koh and B. K. Daniel. Shifting online during covid-19: A systematic review of teaching and learning strategies and their outcomes. *International Journal of Educational Technology in Higher Education*, 19(1):1–23, 2022.
- [18] F. Noorbehbahani, A. Mohammadi, and M. Aminazadeh. A systematic review of research on cheating in online exams from 2010 to 2021. *Education and Information Technologies*, pages 1–48, 2022.
- [19] OpenAI. Chatgpt: A large-scale generative pre-training model for conversation, 2020. URL <https://openai.com/blog/chatgpt/>.
- [20] R. Paul. The state of critical thinking today. *New directions for community colleges*, 2005(130):27–38, 2005.
- [21] R. Paul and L. Elder. *Critical thinking*. Sonoma State University Rohnert Park, CA, 1990.
- [22] K. F. Stanger-Hall. Multiple-choice exams: an obstacle for higher-level thinking in introductory science classes. *CBE—Life Sciences Education*, 11(3):294–306, 2012.
- [23] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention is all you need. *Advances in neural information processing systems*, pages 5998–6008, 2017.
- [24] D. J. Watters and J. J. Watters. Approaches to learning by students in the biological sciences: Implications for teaching. *International Journal of Science Education*, 29(1):19–43, 2007.
- [25] B. L. Whisenhunt, C. L. Cathey, D. L. Hudson, and L. M. Needy. Maximizing learning while minimizing cheating: New evidence and advice for online multiplechoice exams. *Scholarship of Teaching and Learning in Psychology*, 2022.
- [26] A. Y. Zheng, J. K. Lawhorn, T. Lumley, and S. Freeman. Application of bloom's taxonomy debunks the " meat myth". *Science*, 319(5862):414–415, 2008.

A 由 ChatGPT 生成的多项选择题和答案示例

以下是对生成文本研究中考虑的四个学科的多项选择题示例的请求的回应，并提供相关解释的答案。

A.1 机器学习

问题：以下哪一项不是使用支持向量机（SVM）模型执行分类任务的缺点？

- A) SVM 可能对难以选择的核和超参数的选择敏感。
- B) 当数据不是线性可分时，SVM 可能具有较差的泛化性能。
- C) SVM 对输入特征的缩放不敏感。
- D) SVM 在大数据集上训练可能较慢。
- E) 当特征的数量远大于训练示例的数量时，SVM 可能倾向于过拟合。

答案：C) SVM 对输入特征的缩放不敏感。

说明：支持向量机是一种线性分类器，它试图在特征空间中找到最大程度分离类别的超平面。支持向量机的一个缺点是，它们对内核和超参数的选择很敏感，这些参数很难选择，并且会影响模型的性能。当数据不是线性可分时，SVM 也可能具有较差的泛化性能，并且它们在大数据集上训练可能较慢。此外，当特征的数量远大于训练示例的数量时，SVM 可能倾向于过拟合。然而，SVM 对输入特征的缩放不敏感，这意味着即使特征在不同的尺度上，它们也可以表现良好。

A.2 教育

问题：以下哪种理论最好地解释了通过观察和模仿来学习的过程？

- A) Piaget 的认知发展理论
- B) Vygotsky 的社会文化理论
- C) Bandura 的社会学习理论
- D) Bloom 的教育目标分类学

答案：C) Bandura 的社会学习理论

说明：Bandura 的社会学习理论解释了通过观察和模仿来学习的过程，指出个人通过观察和模仿他人的行为来学习。这就是所谓的替代学习或建模。根据该理论，个体不仅通过直接强化来学习新行为，还通过观察和模仿他人的行为来学习新行为。例如，孩子可以通过观看父母示范动作，然后自己练习来学习系鞋带。Piaget 的认知发展理论关注的是儿童的心智发展，以及他们为了理解周围的世界所经历的阶段。Vygotsky 的社会文化理论强调文化和社会因素在学习和发展中的作用。Bloom 的教育目标分类法是一种分类系统，它概述了认知思维和学习目标的不同层次。这些理论都没有具体说明通过观察和模仿来学习的过程。

A.3 市场营销

问题：在以下哪种营销策略中，公司试图通过情感诉求和理想信息来说服消费者购买其产品？

- A) 价格撇脂
- B) 心理定价
- C) 市场细分
- D) 生活方式品牌
- E) 产品定位

答案：D) 生活方式品牌

说明：生活方式品牌是一种营销策略，通过展示公司的产品来提升消费者期望的生活方式或形象，从而与消费者建立情感联系。这一策略通常涉及鼓舞人心的信息传递，并迎合消费者的情感，以说服他们购买产品。其他选项不涉及情感诉求或有抱负的信息，因此它们不是正确答案。

A.4 历史

问题：以下哪一项最好地解释了 1857 年印度兵变对大英帝国的意义？

- A) 它标志着英国在印度殖民统治结束的开始。
- B) 它导致了英国统治的建立，这是英国在印度直接统治的时期。
- C) 它引发了一场关于帝国主义道德的全球辩论。
- D) 它导致东印度公司的废除和权力移交给英国王室。
- E) 它促进了英国工业革命的发展。

答案：D) 它导致了东印度公司的废除和权力移交给英国王室。

说明：1857 年印度兵变，又称印度士兵兵变，是英属印度历史上的重大事件。这是一场反对英国东印度公司的大规模叛乱，而英国东印度公司是当时印度的主要管理机构。这场兵变一开始是抗议使用动物脂肪来润滑步枪子弹，这是对印度教徒和穆斯林的冒犯，但它迅速升级

为反对英国统治的更广泛的起义。兵变最终被镇压，但对大英帝国产生了深远的影响。兵变最重要的结果之一是废除了东印度公司，并将权力移交给英国王室。这标志着印度治理的重大转变，英国政府直接控制了这个国家，建立了英国统治，这一时期的英国直接统治一直持续到 1947 年。