**Z S T U**

**Zhejiang Sci-Tech University**

**计算机学科导论**



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **姓名** | **学号** | | **班级** | |
| **学生** | 李凯涛 | 2023327100056 | | 23计算机科学与技术4班 | |
| **上课时间** | 第3—10周 周四10—11节 | | **电话** | | 15700181194 |

# AIGC

## **一、AIGC生成和AIGC检测技术现状**

**AIGC生成技术现状**

1. 基于深度学习模型（如GANs、Diffusion Models）的内容生成技术广泛应用于图像、文本和音频等领域。
2. 生成技术的快速发展使其可以生成高质量的创意内容，例如自动化图像绘制、小说写作、语音合成、代码生成等。
3. 代表技术包括GPT系列模型、DALL·E、Stable Diffusion等。

某些领域的生成内容质量已经接近甚至超过人类创作者的水平。无论是在新闻内容的撰写、游戏角色的设计，还是在复杂图像和视频的生成方面，AIGC 正在发挥着不可或缺的作用。例如，一些写作助手类工具能依据用户给定的主题、大纲等要求，瞬间生成条理清晰、篇幅可观的文章，极大地改变了传统的内容生产模式。

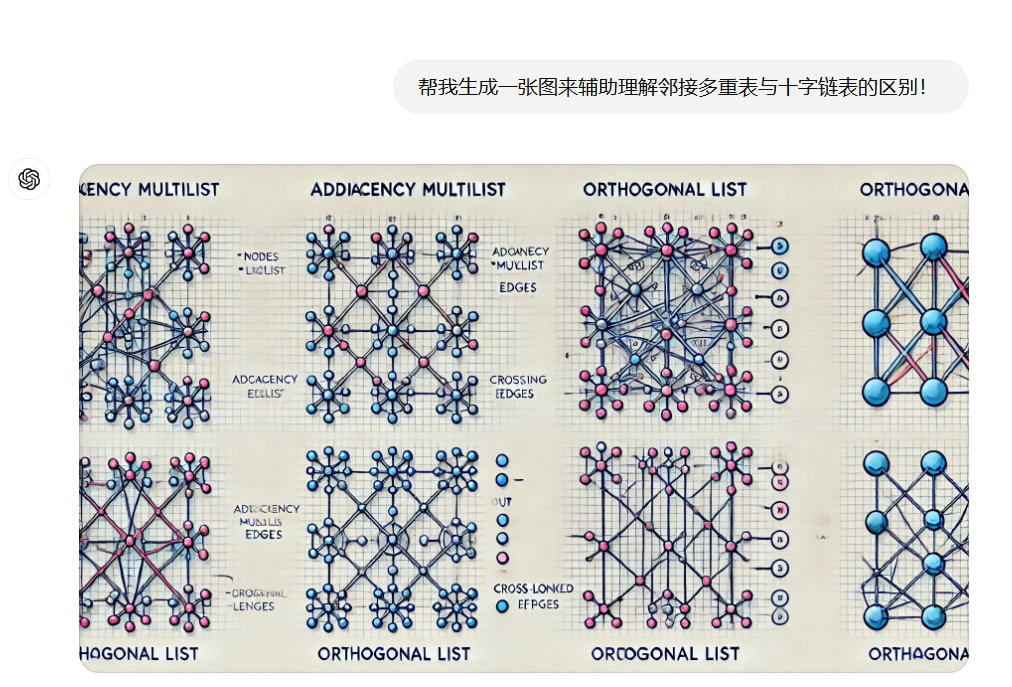


图1：AI辅助学习代码与AI生成图片

 **AIGC检测技术现状**

①随着生成技术的成熟，检测技术逐渐成为重要课题，用于识别生成内容的真实性及其来源。

* 1. 常见应用场景包括学术论文检测、媒体内容真伪鉴定，防止虚假信息传播。

③检测方法正从基于特定模型的特征提取逐步转向通用检测技术，例如基于微观伪影的图像分析和语言模型生成的统计分布检测。



图2：AI对于媒体内容判别不同情况

**二、AIGC生成检测的技术思路**

1. **基于深度学习的图像识别技术**：使用ResNet50模型训练用于分辨真实图像和生成图像的二分类器，并在训练过程中加入了对图像的随机预处理来提升模型的泛化能力。
2. **频域特征分析技术**：经过分析了多种流行的GAN模型生成的虚假图像在频域上的特征，发现这些假图像在频域上都存在着类似的，由上采样操作留下的伪影。
3. **用户行为与内容一致性分析**：结合用户的历史行为和偏好，分析生成内容是否与用户的一贯风格和兴趣相符。通过机器学习算法，建立用户行为与内容一致性的模型，对生成内容进行实时监测和评估，识别出异常或不符合用户风格的内容。
4. **基于语义分析的原创性检测**：通过对生成内容的语义进行深入分析，检测其是否包含独特的见解和深刻的分析。利用自然语言处理技术，对比生成内容与现有文献的语义差异，识别出潜在的抄袭或模板化内容。

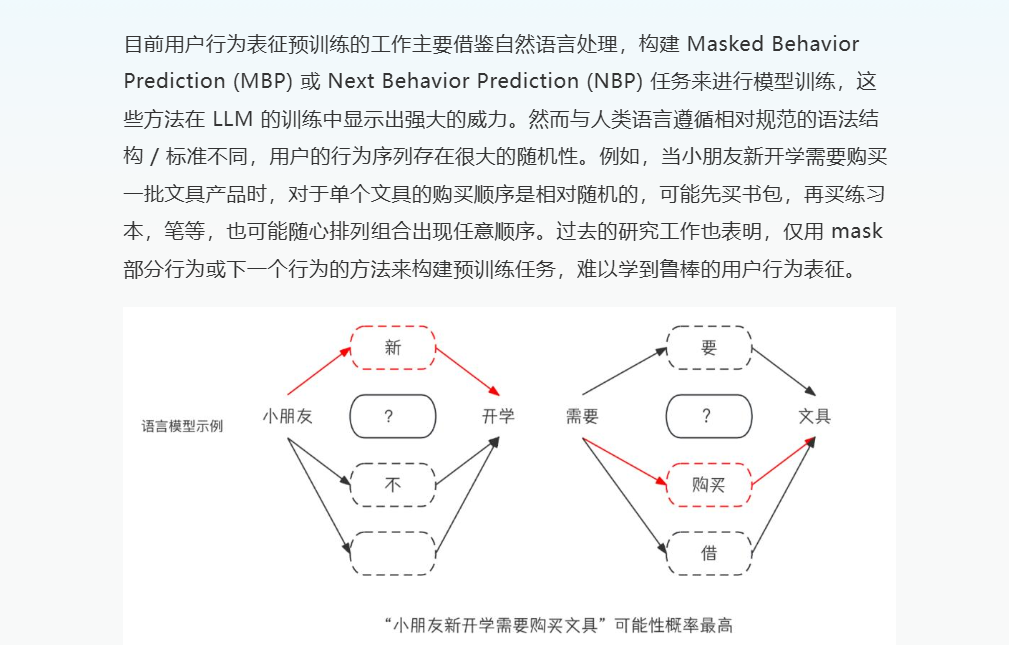


图3：蚂蚁集团构建了Unibehavior - 用户统一行为表征框架三、AIGC对人类发展的影响

**技术角度**：

1. **内容创作的革新**：AIGC技术通过自动化内容生成，极大地提高了内容生产的效率和多样性，降低了创作门槛。
2. **个性化服务的提升**：AIGC技术可以根据用户的需求和偏好，自动生成个性化的内容，如智能推荐、智能客服等，提升用户体验和服务质量。
3. **创新应用的拓展**：AIGC技术的广泛应用将促进创新应用的不断拓展，如在医疗、金融等领域的应用。
4. **推动智能化发展**：AIGC技术的发展将推动各行业的智能化进程，促进数字化转型和产业升级。

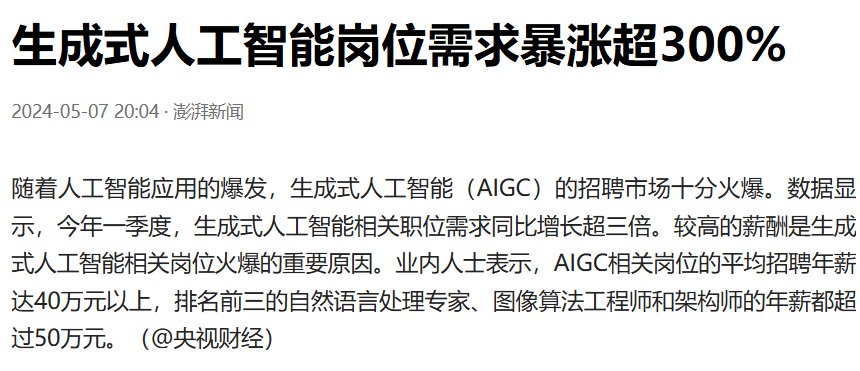
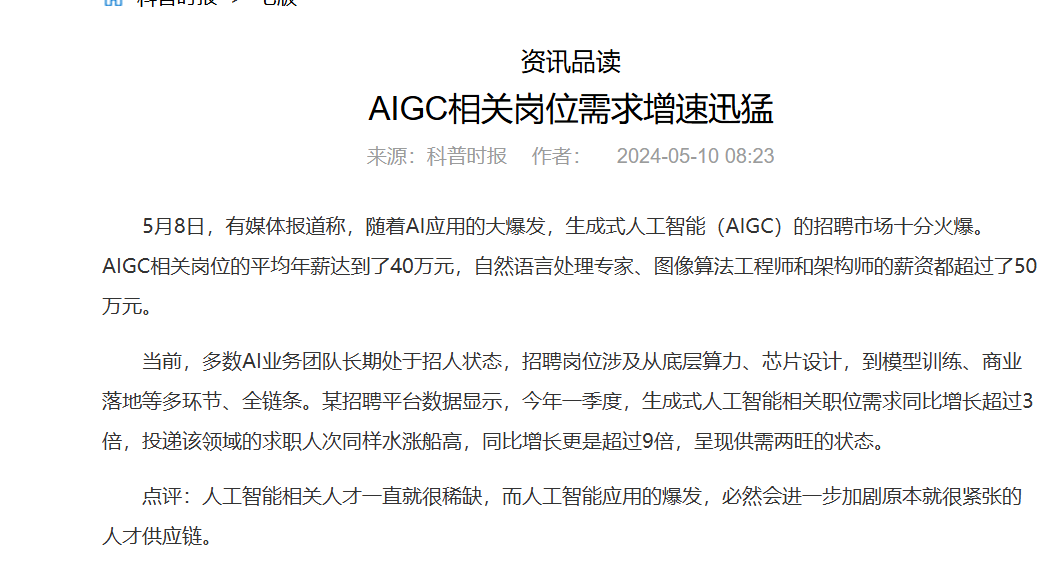
 

图4：AI对就业的提升



图5：AI与医疗领域结合

**合规性角度**：

1. **隐私和伦理问题的挑战**：随着AIGC技术的广泛应用，隐私和伦理问题也将逐渐凸显，如AIGC技术可能会泄露用户的个人信息和偏好，引发隐私泄露的风险。
2. **监管框架的完善**：为了平衡技术发展与社会福利，需要通过反思和调整现有的分配方式，如通过税收和补贴等政策工具来驱动社会向更加平衡和可持续的方向发展。
3. **内容的容错率、AI生成内容的标识等进行规范**：随着《生成式人工智能服务管理暂行办法》的发布，监管框架逐步完善，既展现了对AIGC发展的支持态度，也强调了规范的重要性。
4. **法规与伦理问题**：AIGC技术的快速发展也带来了法规与伦理方面的挑战，如如何制定合适的法规和标准来规范AIGC技术的应用，以及如何处理AI生成内容的版权和伦理问题。
5. **学术诚信问题**：AIGC技术的滥用可能导致学术不端行为的增加，如利用AI生成论文等。因此，需要加强AIGC检测技术的研发和应用，确保学术领域的诚信和原创性。



图6：AI造成的假新闻

**三AIGC对人类发展的影响**

**技术角度**

①提高生产效率：AIGC技术可以快速生成创意内容，降低人力资源成本。

②推动创新发展：广泛应用于文创、教育、科研等领域，为人类提供更多高质量工具。

③丰富内容表达：支持个性化内容定制，提升用户体验。

④低质量内容泛滥：生成内容的过度应用可能降低信息质量。

⑤信息滥用和欺诈风险：虚假内容生成可能威胁社会信任体系。

**合规性角度**

①借助AIGC检测技术，加强知识产权保护和信息来源溯源。

②推动相关法规和行业标准制定，规范技术应用。

③法律责任归属：生成内容的版权归属和违规责任界定仍存争议。

④隐私与伦理问题：模型训练可能涉及个人数据，存在隐私泄露风险。

**四、参考文献**

1、机械之心专栏 作者：蚂蚁机器智能部<https://www.thepaper.cn/newsDetail_forward_25196871>

2、稀土掘金https://juejin.cn/post/7325122758631981093

3、今日头条<https://www.toutiao.com/article/7399485983224021531/?upstream_biz=doubao&source=m_redirect>

4、今日头条<https://www.toutiao.com/article/6498187604150714893/?upstream_biz=doubao&source=m_redirect>

5、澎湃新闻<https://www.thepaper.cn/newsDetail_forward_22548588>

6、中国信通院 人工智能白皮书[P020221111501862950279.pdf](http://www.caict.ac.cn/english/research/whitepapers/202211/P020221111501862950279.pdf)

7、华经情报网<https://www.huaon.com/channel/trend/892504.html>

8、腾讯云<https://cloud.tencent.com/developer/article/2414604>

9、Deepfake Detection Challenge: <https://deepfakedetectionchallenge.ai>.

10、Brown, T., et al. (2020). "Language Models are Few-Shot Learners." NeurIPS.