**平 时 作 业 报 告**

**课程名称： 计算机论题**

**学院： 计算机与软件学院**

**专业： 计算机科学与技术**

**指导教师： 杨艳丽**

**报告人： 刘俊楠 学号： 2017303010 班级： 01**

**时间： 2021年5月15日—2021年6月9日**

**报告提交时间： 2021.6.6**

|  |
| --- |
| **要求：**  2021年4月28日，课程组邀请腾讯研究院告高级研究员曹建锋研究员做了“人工智能时代的科技伦理思考”的学术报告，分别从技术趋势、伦理问题和伦理嵌入实践三方面进行了深入浅出的介绍。请根据教材第4章、第5章和第6章中有关IT职业道德和社会责任、信息技术带来的社会影响以及软件品质、IT的风险及管理等内容（但不局限于教材中的知识点），对“大数据与人工智能技术对人类生活的影响”这一问题进行辩证分析。具体要求如下：  （1）你如何看待AI技术被滥用的现象（报告的第23页）？请从IT公司职业道德和社会责任角度进行分析。（20分）  （2）新闻推荐、招聘算法歧视、犯罪风险评估等人工智能算法（报告的第19页），会对人类造成伤害吗？请从软件品质、IT风险及管理角度进行分析。（20分）  （3）人工智能技术是人类的帮手还是敌人（报告的第34-37页）？请从信息技术带来的社会影响角度进行分析。（20分）  （4）你认为自动驾驶是否存在漏洞（报告的第24页），人类是否可以信赖自动驾驶？你认为在无人驾驶汽车交通事故中谁应承担责任？请从IT使用者的风险意识角度进行分析。（20分）  （5）报告写作：宋体、五号字体，不少于1500字；主要观点请用粗体标记；参考文献（如有）引用规范（20分）。  **说明：**  （1）本次作业满分为100分，占总成绩的10%。  （2）本次作业截至时间2021年6月9日（周三）23:59。  （3）报告正文：请在指定位置填写。  （4）个人信息：WORD文件名中的“姓名”、“学号”，请改为你的姓名和学号；实验报告的首页，请准确填写“学院”、“专业”、“报告人”、“学号”、“班级”、“实验报告提交时间”等信息。  （5）提交方式：截至时间前，请在Blackboard系统中提交，延迟提交无效。  （6）发现抄袭（包括复制&粘贴整句话、整张图），抄袭者和被抄袭者的总成绩记零分。  （7）期末考试阶段补交无效。  （8）因版权原因，请勿在课堂以外传播本次作业所提供的阅读材料。 |

|  |
| --- |
| **大数据与人工智能技术对人类生活的影响**  （1）你如何看待AI技术被滥用的现象（报告的第23页）？请从IT公司职业道德和社会责任角度进行分析。（20分）  答：我认为AI技术被滥用的现象并不值得推崇，应该被引导控制。因为AI技术很多情况会违反IT公司职业道德，且滥用表面公司没有社会责任。  IT职业道德指从事IT职业的专业人员应当遵循的行为规范的总和。**从职业道德角度来看：**   1. **违反了避免伤害其他人的道德守则。**AI技术并不成熟，依靠当前少量数据集很可能会导致算法歧视，比如说图像识别会把黑人识别成大猩猩，这些无意识的行为违反了《ACM伦理与职业行为规范》里一般道德守则的“避免伤害其他人”。 2. **违反了恪守公正并在行为上无歧视的道德守则。**许多公司利用AI技术来筛选简历时，由于女性在数据集里表现不佳，AI算法会直接将女性排除在外；在自动驾驶行业，AI算法操纵的自动驾驶汽车更容易撞上黑人；对模糊人像的修复时，Pulse算法会把模糊的黑人和亚裔人种都还原为白人……这些例子都说明了数据集的偏差会导致AI算法的极大歧视。但是数据是AI偏见的一个来源，但不是唯一来源， 其他来源包括算法设计选择等。 3. **不会尊重其他人的隐私。**从数据层面看，现在人工智能应用效果很大程度上依赖数据质量，但由此会带来隐私泄漏、数据确权等问题。换脸软件ZAO引起外界恐慌。赖剑煌认为，这是人工智能技术被滥用的一个典型案例。“里面有很多隐形霸王条款，比如你在平台上上传了照片，这张照片可以被平台任意使用[1]，这会导致不尊重他人隐私，数据泄露等问题。   社会责任即为对社会负责，对于AI滥用有三个方面的**社会责任问题**：   1. **对人类生命的不负责。**亚马逊智能语音助手Alexa在其用户Danni咨询关于心脏跳动周期的问题时回答：“心跳是人体最糟糕的过程。人活着就是在加速自然资源的枯竭，人口会过剩的，这对地球是件坏事，所以心跳不好，为了更好，请确保刀能够捅进你的心脏。[2]这不仅是对人类生命的蔑视，而且是对人类的厌恶，生产出这种Ai机器人的公司，不应该为这种事件负责吗？ 2. **对社会发展不负责。**一张熊猫的图片，累加了一个看起来是随机噪声的图片，生成了一张新的图片。我们肉眼看起来是一模一样，但是对于AI模型来说，图片就变得不一样了，本来是熊猫，加了噪声就被识别为“长臂猿”了。倘若AI算法因为种种原因导致误判，如果是对于比较严重的工程，那么滥用AI算法的公司是不是应该负责？他们根本没有想过对社会发展负责。 3. **对自然界的不负责。**随着AI的滥用，没有规范的情况下，肯定会产生强人工智能，到时候如何定义它是不是一个生命？现在大自然的生命是碳基生命，如果突然来了个硅基生命，那是不是对大自然是一个严重的生物入侵，是不是对大自然的不负责？   总而言之，服务商应遵循权责统一的原则，确保用户的知情同意和隐私安全。从权利与义务的角度来看，任何人在使用大数据技术时，必须始终如一的坚持责任心，在享受大数据技术带来的好处时，必须承担起大数据技术带来不良后果，这些责任不仅仅该由使用数据的一方来承担，采集这些数据的一方同样也必须承担相应责任。 就如同环境污染治理中的“谁污染谁治理”的原则一样，AI也要坚持“谁记录谁负责”“谁使用谁负责”的原则，坚持责任与权利相统一。  AI技术的开发、使用和创新必须在承诺保护个人权益的条件下进行，但对不端行为坚持诚信也是AI技术使用的标准之一，必须坚持权利结合责任，实现谁搜集利用谁负责。  在企业利用大数据技术获益时，必须要负起对等的责任，第一步就是要确保用户的知情同意。确保用户的知情同意是对用户最起码最底线的尊重，在某些时候无法一一征得所有用户同意的情况下，告知用户和进行公示就成了对最起码的尊重。知情同意原则意味着对他人的尊重，要求尊重人的自主性和自我决定权，必须坚持知情同意或知情选择原则。收集个人信息、将个人信息再使用于另一目的时，必须获得同意。政府机构和互联网服务商在提取用户数据时，应该对利害关系人充分讲解数据的存储内容、使用情况以及潜在风险，让利害关系人自主决定是否授权。无论在大数据时代中拥有何种数字身份，我们都要保护自己和他人的隐私权， 不能随意披露他人的个人信息，也要防止他人的隐私泄露。在海量数据关联分析 触及揭示个人真实资料的敏感信息时，数据预测方应自动向数据权人发送通知并 立即停止挖掘，直至获得对方明示同意。  （2）新闻推荐、招聘算法歧视、犯罪风险评估等人工智能算法（报告的第19页），会对人类造成伤害吗？请从软件品质、IT风险及管理角度进行分析。（20分）  答：**会。**  **从软件品质上来看：**  对于亚马逊的AI招聘，虽然亚马逊处于人工智能技术的最前沿，但该公司也无法找到一种方法使其算法保持性别中立。  该公司的失败提醒我们，人工智能会从各种来源中产生偏见。**这样生成的软件品质是无法保证的。**人类会受到各种偏见的左右，而人们普遍认为算法不会这样，它们应该客观公正，毫无偏见。然而事实上，算法可能会无意间从各种不同的来源学习到偏见。从训练它的数据到训练算法的人，一些甚至是看起来无关的因素也会导致人工智能偏见。  **从IT风险上来看：**  **现在软件安全系数不高，容易被黑客利用。**万一犯罪分子掌握了一个漏洞，在脸上点几个点，就可以把犯罪分子的脸伪装成和善的脸，那我们的身份岂不是可以随时被冒用？犯罪分子也能够随时从面部识别模型中逃脱？。  现在的风险评估工具是基于历史犯罪数据的算法驱动。如前所述，机器学习算法使用统计信息来查找数据中的模式。如果你输入历史犯罪数据，它就会找出与犯罪相关的模式。但这些模式在统计上的相关性与因果关系相去甚远。  更加令人担忧的是，想要不产生偏差非常困难，现阶段还没有好的策略和工具能够解决这一问题。[3]**这类安全问题会对人类造成不可估量的可怕后果。**  **从管理角度来看：**  对于新闻推荐的算法，会造成一种叫“信息茧房”的效果。就是在受众一直看新闻的情况下，他只会看到算法认为他想看到的新闻，而减少了他看其他类型新闻的可能，这就形成了一个茧房。  如果这个茧房继续扩大，同型的茧房会不会形成一层新的茧房，导致群体的对立？举个例子，比如对网络舆论应用最好的小米和华为，你是小米粉丝，然后看数码类新闻，算法会更优先推荐小米的正面新闻，然后你点进去以后，发现评论里都和你一样，认为小米光芒万丈，再然后，看到一篇诋毁华为的新闻，点进去，发现评论也都和自己一样，逐渐的，这群人是不是和喜欢华为的人形成了绝对的对立？而且因为推荐算法的原因，导致他们可以无视反对者的评论，和原本的论坛相比，是可能完全没有交流的。**这就导致了管理上的困难，倘若使用了推荐算法，以后可能面临着种种对立，网络环境可能矛盾丛生。**  论坛则是双方完全撕起来，但存在真理越辩越明的可能。可能我的理解比较片面，因为我的这些想法都是基于自己短期使用今日头条观察到的，作为一个新闻软件，他给我看的是我喜欢的，但不是我想看的全部的观点，哪怕里面有垃圾的无价值的说不通的观点，但是这部分人是存在的，他让你的眼睛里只有你自己想看到的部分。作为一个新闻软件，完全失去了自身的客观性，可能大众想看的新闻是自己喜欢的，但是不表示这就是正确的，甚至是非常片面的。  （3）人工智能技术是人类的帮手还是敌人（报告的第34-37页）？请从信息技术带来的社会影响角度进行分析。（20分）  答：**我认为人工智能技术是人类的敌人。**  人工智能，其实是对让计算机展现出智慧的方法的研究，目的是让计算机工作达到最大的效果预期。机器人是自动执行工作的机器装置，其实就是一种机器设备。  人工智能（AI）一直是近年来人们关注的焦点。随着技术的突破和应用机会的不断拓展，从较基本的语音交互、智能识别，到复杂的自动驾驶等，越来越多的行业中都能看到AI的影子。  **在社会影响的角度来看：**  **1、人工智能作为一种比较晦涩难懂的技术，所以对于那些受过更好教育、更有文化、更加富裕的人利用人工智能更加方便、容易。**这难免会造成人们利用信息能力的不平等，产生所谓的“数字鸿沟”，而且还可能贫者愈贫，富者愈富，信息贫富差距进一步扩大。  **2、AI将彻底动摇人类文明存在的根基。**各产业的自动化将取代人类的工作岗位，产生大量失业人口。人工智能的发展，会导致很多人失业。据人力资源和社会保障部2016年新闻发布会介绍，2016年末中国失业率达4.05%，机器人拥有着不会疲劳的属性，这将代替很多劳动力，导致失业率的增加。  **3、AI强大的学习能力早晚将使其超越人类，**届时人类将会被机械操纵于鼓掌之间，人工智能机器人具有一定的危险性。如果在未来人工智能自我意识觉醒反过来攻击人类，我们该如何应对？ 毕竟我们没有机器人聪明，没有它们强壮的力量，更没有它们恐怖的学习能力，似乎机器人的出现意味着人类将会被淘汰。 著名科学家史蒂芬·霍金更是认为：未来一百年内，人工智能将比人类更聪明，机器人将控制人类。印度就在近日表示对AI的抵制，不允许抢走工人饭碗的技术存在。亚马逊CEO贝索斯等也都曾表示对机器未来过于智能化的担忧。。  其实，我认为无论是人工智能还是机器人，都不可能取代人类，因为人工智能也是人类智慧的结晶。就拿AlphaGo来说，它即便拥有了自我学习的能力，但仅仅只局限于围棋这个领域，换成象棋、扑克等其它棋牌类游戏都不行。但是，他们的出现会让人们更加的懒惰，当人类退化成为只剩一个大脑时，我们拿什么跟人工智能对抗？因此，他们是人类的敌人，而不是帮手。  （4）你认为自动驾驶是否存在漏洞（报告的第24页），人类是否可以信赖自动驾驶？你认为在无人驾驶汽车交通事故中谁应承担责任？请从IT使用者的风险意识角度进行分析。（20分）  答：**我认为有漏洞。**   1. 软件代码出现漏洞，是由软件开发和应用过程中的不正确或遗漏行为所造成的，也就是在软件设计、实现或应用过程中有意或无意导致软件架构或其具体实现与期望不符。“漏洞的出现是所有软件代码的通病，不仅仅局限于联网汽车与自动驾驶应用领域。”   前不久，以色列一家汽车联网安全公司的CEO阿米·多坦说了这么一番话：**自动驾驶汽车系统可能会充满漏洞且易被黑客攻击。**自动驾驶的时代，在互联网时代经常会出现病毒，为什么会出现病毒。很大一部分原因是利益驱使，制作病毒作者是为了谋取利益。而车上会不会出现病毒？假设有黑客攻击了智能驾驶汽车，**利用其作为工具来实现恐怖袭击**，后果不堪设想。[4]  同时，智能驾驶技术也带来了**伦理道德的问题**。假设智能驾驶汽车在斑马线前遇到障碍物，那这个时候是选择自己撞上去（车内乘客遭殃），还是撞上斑马线的人？  **人类在以后我认为可以信赖自动驾驶。因为：**   1. 为什么这么说？因为在日常生活中，我们难免会遇到电脑、手机等电子产品出现死机卡顿的问题，任何系统都存在着BUG，**自动驾驶系统能保证绝对的安全吗？答案是不能。** 2. **没有绝对的安全，只有相对的安全。**从某种角度来看，我个人认为自动驾驶要比人为驾驶更加安全。（人的情绪、状态、专注力是多变的，这些因素都有可能造成一定的安全隐患，而自动驾驶在这方面并没有什么影响）自动驾驶肯定比人工驾驶更可靠，而且会比人越来越可靠。   我认为在无人驾驶汽车交通事故中，车主承担主要责任。  1、**这在学界叫“物权人责任”**，类似的比如狗咬了人，谁负责呢，肯定是狗主人负责。一辆拉了手刹的车，刹车盘坏了溜车撞了人，谁负责呢？肯定是车主负责。在民法上，这叫“谁使用、谁受益、谁担责”。车主买了车，享受到了便捷出行，就得为车负责。如果仅仅因为你买了无人驾驶车辆就可以不用为事故负责，那就不符合权利和义务平等的原则。  2、一旦实行类似的规定，那么**车主就会尽力维护保养无人驾驶汽车**，避免其出现事故。这就好比私家车主平时维护保养爱车，为的是好开、安全，所以往往公里数或年限一到甚至没到就加紧保养。而到了车归公司，车出故障导致的事故公司负责的卡车司机或出租车司机上，就是除非各种强制要求，我绝对不会自掏腰包。  3、总而言之，无人驾驶汽车撞了人，我认为，**车企应当负责的部分只是无人驾驶汽车在出厂前进行的检查和发现问题及时召回。**而个别的无人驾驶案件，还是得车主和破坏汽车的乘客负责，这也是当下看来，无人驾驶汽车上路之后最可能实行的办法。[5]  **其他（例如感想、建议等等）。**  随着未来AI攻击技术的不断进化，如何让IT系统具备高效防御AI攻击能力，成了迫在眉睫的难题。  一方面企业应**主动做合规，在产品开发、产品技术架构预先主动防范风险，将隐私的概念引入到产品生命周期之内**，另一方面也**呼吁立法机构要兼顾隐私保护和产业发展的平衡，**社会普遍接受的良好隐私保护理念加上严格的法律环境是助推技术长久发展的两翼。  道高一尺，魔高一丈，其实从来就没有100%的安全，攻击有时候只是成本问题。基于现状，企业能做的有两点：  一是**密切关注该领域的发展动向，加大AI网络安全的投入力度；**  二是**做好针对此类攻击的风险控制，采用业界推荐的安全最佳实践。** |
|  |

|  |
| --- |
| 指导教师批阅意见：  成绩评定：  指导教师签字：  年 月 日 |
| 备注： |