

人工智能推荐系统中的大数据杀熟

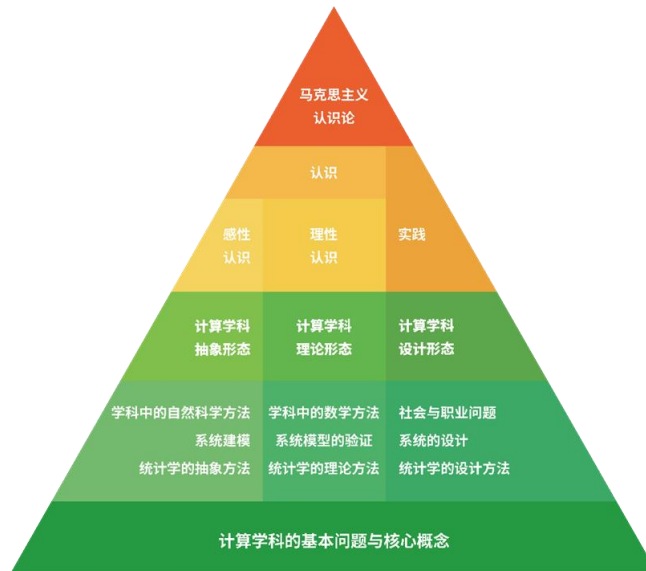
计算机课程思政虚拟教研室 工程伦理样例 1

一、教学目标

本案例能够置于 Bloom 分类法知识维度的“元认知知识”位置，学生学习后能够达到 Bloom 分类法认知过程维度的“评估”层次。

二、本案例课程思政的关注点

1. 本案例属于学科设计形态中的社会与职业问题。算法歧视是由数据和算法分析导致的对特定群体的，系统的、可重复的、不公正的对待。平等与自由共同构成现代法治的基础，公正是现代文明国家的价值追求，算法歧视令特定群体遭遇系统性的不公正对待，在起点上违背平等的理念，在终点上导致不公正的后果，从根本上违背平等和公正的社会主义核心价值观。



2. 在本案例中，要求教师将 CC2020 报告中 11 个品行元素中的“主动性、责任感、专业性”，CS2023“人工智能”中的“严谨性、坚持不懈”品行，以及“社会、伦理与职业化”中的“批判性自我反思、反应灵敏、文化胜任力、倡导”品行对齐，并与该案例绑定在一起进行可操作性解释。

三、本案例中的抽象、理论和设计三形态

“大数据杀熟”是推荐系统中存在的一种典型的偏见歧视方式，是指企业分析和预测消费者大数据，对不同对象收取不同价格的行为，尤其是，对于同样商品或服务，老客户得到的价格，反而高于新客户，亦即，老客户被“杀熟”。本案例归属于计算学科设计形态的内容。

目前“大数据杀熟”已呈泛化趋势，不局限于对熟客的“宰杀”。经营者通过大数据分析 with 用户画像，可以针对任何其所认为的“对价格不敏感”的用户实施差别定价。互联网企业掌握了大量个人信息、行为习惯等数据后，判断其偏好、用户黏合度、价格敏感度等，利

用大数据技术实现“千人千面”，使不同用户看到不同的价格或搜索到不同的结果，从而被“杀熟”，这也被称为“杀熟二代”。杀熟二代中的“熟”，已经不是“熟客”，而是被平台充分掌握个人信息的“熟人”。“大数据杀熟”一般是指交易价格的差别待遇，因此“大数据杀熟”一般被定性为“差别定价”或“价格歧视”。

价格歧视（Price Discrimination）实质上是一种价格差异，通常指商品或服务的提供者在向不同的接受者提供相同等级、相同质量的商品或服务时，在接受者之间实行不同的销售价格或收费标准。经营者没有正当理由，就同一种商品或者服务，对若干买主实行不同的售价，则构成价格歧视行为。

根据差别定价的程度，英国经济学家庇古（Arthur Cecil Pigou）把价格歧视区分为三个等级：

（1）一级价格歧视（First-degree Price Discrimination）又称为完全价格歧视，即，销售者为每一位顾客及其所购买的每一单位商品制定不同的价格，因此获取所有的消费者剩余。一级价格歧视比较少见。

（2）二级价格歧视（Second-degree Price Discrimination），即，销售者对购买者偏好的多样性有所了解，但是不能观察到每一位特定顾客的特性。即便如此，仍然有可能通过提供一系列包括价格和各种条款的销售合同，在不同的购买者之间进行价格歧视。

（3）三级价格歧视（Third-degree Price Discrimination）又称为作指示选择（Selection by Indicators），即，有时企业拥有关于顾客的信息，特别是关于顾客愿意支付的价格的信息。如果购买者的特性是可以观察到的，那么销售者就可以制定不同的价格，此时，价格就可以看作购买者特性的函数。三级价格歧视是最普遍的价格歧视形式。三级价格歧视中销售者将购买者分组，对每一组制定不同的价格，这种行为也称为市场分割（Market Segmentation）。

1.滴滴“苹果税”

2020年，复旦大学管理学院副教授孙某带领研究团队在五座城市打了800多趟车，花费50000多元，搜集到滴滴、曹操、首汽、美团和高德等多个渠道的数据。调查招募了20多名大学生作为调研员，对五城市（上海、北京、深圳、成都和重庆）、不同距离（近途：3公里以内，中途：3-10公里，远途：10公里以上）以及多个时间段（工作日早高峰、工作日晚高峰、日间非高峰、晚间非高峰）进行了抽样调查。获得总样本836个，有效样本821个，通过数据分析，最终形成“2020打车报告”。

该研究团队表示，其在研究过程中验证了“苹果税”的存在。团队用“一键呼叫‘经济型+舒适型’两档后被舒适型车辆接走的订单比”来判断“被舒适”的程度，数据表明，与非苹果手机用户相比，苹果手机用户更容易被舒适型车辆（如，专车、优享等）司机接单，这一比例是非苹果手机用户的3倍。

除了手机品牌，平台也可能同时关注乘客手机价格。如果乘客使用的是苹果手机，那么就更容易被推荐舒适型车辆；如果乘客不是用苹果手机，那么就要看其手机价位，手机价位越高则越有可能被舒适型车辆接走。

此外，“苹果税”还体现在打车优惠上。数据表明，苹果手机用户平均只能获得2.07元的优惠，显著低于非苹果用户的4.12元。除绝对金额外，优惠折扣比同样支持上述结论。

早在2018年，针对滴滴大数据杀熟的质疑，滴滴总裁柳青在微博澄清表示，因为路况是动态变化的，前一秒和后一秒都可能发生很大变化。为了给乘客尽可能准确的预估参考价，滴滴是实时根据动态路况预测预估价的，即便起终点一样，预估价也会存在差异。并且，两个用户的预估路径规划可能不同，对经常打某一路线的用户会用他平时常用的路径缺省值，没走过这个路线的用户系统会推荐最优方案。

2.携程“钻石会员”服务

胡女士是携程的“钻石会员”，2020年7月，胡女士通过携程APP订购了舟山希尔顿酒店的一间豪华湖景大床房，支付价款2889元。离开酒店时，胡女士发现酒店的实际挂牌价仅为1377.63元。

胡女士认为，自己不仅没有享受到星级客户的优惠，反而多支付了一倍的房价。退房后，她与携程沟通，携程以其系平台方，并非涉案订单的合同相对方等为由，仅退还了部分差价。胡女士不服携程的处理结果，以携程公司采集其个人非必要信息，进行“大数据杀熟”等为由起诉到法院，要求“退一赔三和携程APP为她增加不同意《服务协议》和《隐私政策》时仍可继续使用的选项，以避免携程公司采集她个人信息，掌握她的数据。”

绍兴柯桥法院审理此案后认为：携程APP作为中介平台，对酒店实时房价有如实报告的义务，但其却未如实报告。携程向原告胡女士承诺钻石贵宾享有优惠价，事实上却没有价格监管措施，向原告展现了一个溢价100%的失实价格，未践行承诺。携程在处理原告投诉时告知原告无法退还全部差价的理由，经调查也与事实不符，存在欺骗。

法院认定，被告携程公司存在虚假宣传、价格欺诈和欺骗行为，支持原告“退一赔三”的诉讼请求。而且，携程APP收集的用户信息超越了形成订单必需的要害信息，属于非必要信息的采集和使用，其中用户信息分享给携程可随意界定的关联公司、业务合作伙伴进行进一步商业利用更是既无必要性，又无限增加用户个人信息使用的风险。

据此，法院当庭作出宣判，判决被告携程公司赔偿原告胡女士投诉后携程未完全赔付的差价243.37元及订房差价1511.37元的三倍，总计支付赔偿金4777.48元，且应在其运营的携程旅行APP中为原告增加不同意其现有《服务协议》和《隐私政策》仍可继续使用的选项，或者为原告修订携程旅行APP的《服务协议》和《隐私政策》，去除对用户非必要信息采集和使用的相关内容，修订版本需经法院审定同意。

2021年7月19日，携程就此次判决回应21世纪经济报道记者，胡女士遭遇的价格差系供应商囤房再利用库存紧张的时间点进行超高价销售，并非“价格杀熟”。此外，携程在对外情况说明中提到，公司已经认识到仍需强化对供应商“囤房加价”行为的监管拦截，并上线新版风险控制系统。

3.美团“割韭菜”

2020年12月，一篇名为《我被美团会员割了韭菜》的自媒体文章成为全网爆款，再次把大数据杀熟的问题推上风口浪尖，美团成为众矢之的。名叫“漂移神父”的网友爆料称，他在美团App购买驴肉火烧外卖时发现，在同一家店铺、同一个配送位置、同一个下单时间点的条件下，开通美团会员的账号所需的配送费，比未开通会员的账号所需的配送费高出整整4元。该网友多次亲身测试后指出，其所在位置附近几乎所有外卖商户，会员账号总是超出非会员账号1-5元不等的配送费。

在反馈客服后，对方表示愿意补偿十元红包。电话客服主管则表示已超出他的处理范围，会向更高级管理人员反映。文章中还提到，不仅是一家店有这种情况，一部开通美团外卖会员的手机，附近几乎所有外卖商户的配送费，基本都要超出非会员配送费1-5元不等。换句话说，会员并未享受到平台所承诺的优惠待遇，只要多点几次外卖，额外支付的配送费甚至会超过会员所享的优惠。

针对用户反馈的“会员和非会员配送费差异”问题，美团外卖公开回应称，文中提到的配送费差异与会员身份无关，是软件的定位缓存偏差导致。用户实际下单的配送费，会按照真实配送地址准确计算，不受影响。

“美团的解释我并不认同。”文章作者在接受媒体采访时表示，第一次地址是有问题，但第二次地址是一样的，因此他不认同美团的回应。此后，文章作者再发微博表示，本次事件已然会改变大家对互联网巨头的认知，也可能会促成巨头更合规的运营。

四、专业品行

将大数据杀熟案例与 CC2020 报告中的“主动性、责任感、专业性”，CS2023 “人工智能”和“社会、伦理与职业化”领域的品行要求（“严谨性、坚持不懈”和“批判性自我反思、反应灵敏、文化胜任力、倡导”）绑定在一起，可以为推荐系统的设计和 implement 提供明确的伦理指导和操作框架。

1.主动性 (Proactivity)

在大数据杀熟背景下，主动性要求学习者不仅了解大数据杀熟的表现形式，更要主动寻找并解决可能导致消费者不公平对待的系统内部问题。例如，主动评估推荐算法，给出合理的修改方案，确保消费者不会因购买历史或使用的设备类型而导致被价格歧视。

2.责任感(Responsible)

面对大数据杀熟问题，责任感意味着对消费者权益的保护承担起责任。这包括确保推荐系统的透明度，让用户明白他们的数据如何被使用，以及为可能的不公平定价行为提供补救措施。

3.专业性(Professional)

专业性在大数据杀熟场景中体现为使用专业知识主动识别和防止可能的歧视行为。这意味着学习者需要深入了解当前技术可能造成的社会影响，并开发出既高效又公正的算法。

4.严谨性(Meticulous)

严谨性要求学习者在设计和实现推荐算法时，对数据的处理、使用和结果的影响进行严格考查，避免数据采集、处理和分析阶段可能引入的偏差。

5.坚持不懈(Persistent)

对于坚持不懈，在遭遇算法偏见和不公平的定价问题时，学习者应不断研究数据集的特征，持续改进算法，不断测试和调整，直到找到既能满足业务需求又能保护消费者权益的解决方案。

6.批判性自我反思(Critical Self-reflection)

批判性自我反思在该案例中意味着学习者开发的项目出现歧视时，应该及时反思技术和业务实践，尽早识别那些可能导致或加剧不公正的行为，并采取行动纠正这些问题。

7.反应灵敏(Responsiveness)

在应对大数据杀熟问题时，反应灵敏要求学习者作为公司的技术人员对外界的批评和消费者的反馈迅速响应，及时调整其业务和技术策略以解决这些问题。

8.文化胜任力(Cultural Competence)

文化胜任力强调学习者需要理解不同消费者群体的特性和需求，确保推荐系统的设计和运营不会无意中忽视或歧视某些群体。

9.倡导(Advocacy)

倡导在大数据杀熟案例中意味着学习者不仅在自己的工作中实践公平原则，还要在更广泛的社会和专业层面推广该原则，包括参与公共讨论，推动制定和改进相关法律法规。

在大数据时代，学习者具有和实践上述品行，不仅可以提升推荐系统的公平性和透明度，还能增强消费者对平台的信任，最终促进企业的长期发展和社会责任的实现。

五、激励、唤醒和鼓励同学们向上的途径

数据本身没有价值倾向，但“大数据杀熟”明显走了歪路，对消费者、社会将产生长期危害。对于消费者个人来说，一次交易中或许仅仅是价格的小数额差异，但当消费者对平台产生依赖，长期而言，交易成本不降反升，侵害了消费者的公平交易权；对于整个社会而言，杀熟行为破坏了诚实信用的基本原则，降低了消费者的消费欲望，导致交易成本的上升。该

案例唤醒同学们的人工智能伦理意识和社会责任意识,激励同学们在未来的职业生涯中自觉抵制“大数据杀熟”类似行为,自觉遵守人工智能伦理规范,促进互联网平台良性发展。

六、习题

- 1.根据文档中的定义,解释什么是“大数据杀熟”以及它在推荐系统中的表现。请列举至少两个实际案例来支持你的解释。
- 2.讨论“大数据杀熟”对消费者和社会的潜在影响。请考虑但不限于以下方面:消费者权利、市场公平性、社会信任和企业道德。
- 3.根据案例描述,讨论造成算法歧视的原因和后果,并提出可能的解决策略。
- 4.基于文档中提到的携程“钻石会员”服务的案例,分析该案例中的问题点,探讨携程如何通过改善其服务或政策来避免此类“大数据杀熟”的情况。

参考文献

- [1] 古天龙.人工智能伦理导论[M].高等教育出版社,2022
- [2] 古天龙,李龙,常亮,罗义琴.公平机器学习:概念、分析与设计[J].计算机学报,2022,45(05):1018-1051
- [3] Xuan Feng, Tianlong Gu, Xuguang Bao, Long Li. Behavior-based Ethical Understanding in Chinese Social News. IEEE Transactions on Affective Computing, doi:10.1109/ TAFFC. 2022. 3160745
- [4] 常亮,张伟涛,古天龙,孙文平,宾辰忠.知识图谱的推荐系统综述[J].智能系统学报,2019,14(02):207-216
- [5] 董荣胜,古天龙,殷建平.计算学科课程思政教学指南 [J]. 计算机教育,2024(01): 7-15.