

客户满意的深度洞察

——结合企业内外部 UGC 的可解释深度学习

2021 年获国家自然科学基金项目：基于企业内外部用户生成内容匹配和可解释深度学习技术的客户满意度分析
项目负责人：清华大学经济管理学院管理科学与工程系 卫强教授

客户满意是影响客户购买、留存、忠诚以及企业口碑的重要决定因素，与公司的长期利润与价值创造密切相关。随着社会化平台的不断推广应用，用户生成内容（User Generated Content, UGC）得到企业的广泛关注。特别是客户评论这类企业外 UGC，以其自主性、丰富性和多样性，成为企业分析客户满意的重要信息源。同时，由于客户评论的大数据特征（非结构化、数量巨大、低价值密度等），相关研究设计了多种机器学习方法对其进行特征提取、情感极性探测、以及关键影响因素挖掘等。近几年，深度学习方法以其强大的学习能力和高水平的预测精度，也在客户满意分析领域开始得到关注和应用。

但现有基于 UGC 进行深度学习并支持客户满意分析的研究还存在如下挑战。第一，客户评论具有“强情感，弱内容，自发性”的特点，使得其蕴含的信息质量较差，因而难以实现对影响客户满意的因素和缘由的深度洞察。这一挑战在高专业化程度的服务场景下更为突出。第二，大量分析方法直接将机器学习模型应用于客户评论文本数据，缺乏适当的管理理论和框架的有效支撑和合理集成。易使挖掘得到的特征或因素比较琐碎、零散和表面化，且缺乏可靠的管理意义，从而造成实践指导意义较弱。第三，深度学习模型的引入虽有助于提升学习精度，但由于其复杂的“黑箱”机制，如繁复的序列规则、成百上千的隐藏状态单元和数百万个参数的更新等，使得所学习到的模式的管理可解释性不足，无法真正用于支持业务改进和服务优化。



针对上述挑战，本项目首先引入另一类重要的信息源，即企业内 UGC——企业 CRM/内部博客中的员工日志等。相对于客户评论，员工日志是由企业员工针对特定客户业务所记录的信息，如投诉细节、问题症结、潜在解决方案、处理进度和客户状态等信息，具有“弱情绪，强内容，规范性”的特点，可以极大弥补客户评论的不足。更重要的是，员工日志与客户评论是分别从企业内外部不同视角对于同一业务的描述，这样的双边信息源的匹配为深度洞察客户满意度提供了有力支撑。

其次，通过对客户满意测度和服务质量分析的相关文献和理论的梳理发现，客户的满意度受到情感和认知两个组件的共同作用。其中情感组件反映的是客户自身的态度，而认知组件则取决于客户对服务质量的期望与企业提供的服务质量之间的感知差距，即可提炼出情感-认知服务质量差距理论模型。传统上，此模型更多被作为概念模型使用，并没有与机器学习/深度学习进行整合。而如能将此具有管理内涵的概念模型与深度学习方法

进行有机结合，则可以在保留高精度学习性能的前提下，打开该模型的部分“黑箱”，提升模型和学习结果的管理可解释性。

基于上述思路，本项目创新性地将客户满意度分析的相关理论框架，与深度学习的技术框架进行交叉融合，并整合企业内外部双边 UGC，进行建模和分析，从而形成对客户满意的深度洞察。特别是基于情感-认知服务质量差距这一影响客户满意的理论模型，构建相应的内生作用机制，通过深度学习将该理论模型进行合理建模和嵌入，使得所设计模型兼具高水平的预测性能和较好的可解释性这两方面的优势。进一步，项目还将基于所搭建的模型和方法框架，设计并计算“归因指数”，以获取深度模型中间层的潜在管理语义，为企业提供新的可改进的影响客户满意度的潜在因素，并能为有效支持管理诊断分析。

本项目的研究工作正在稳步推进中，并已取得了一些初步成果。特别是项目团队与某大型电信企业合作，基于其客服场景，将所得到的理论模型和方法进行应用和示范，取得了良好效果。所设计的满意度因素分析与诊断原型系统 Deep NPS Analytics 得到该企业的重视，在内部战略报告里作为关注热点进行了报道，并指出该方法有效提升客服部门绩效的重要技术趋势之一。

此外值得强调的是，相对于传统工程场景，管理场景具有跨域高维，判别标准模糊多变以及受人的主观性影响等挑战，使得应用深度学习技术来解决管理决策问题会面临更凸显的“黑箱”现象和可解释性不足的问题。本项目针对这一重要的研究方向，做出了有益探索。特别是所设计的将具有内生可解释性的管理理论框架与深度学习模型进行有机集成的技术路线，可在很大程度上兼顾高学习精度和良好的管理可解释性，具有重要的泛化意义。

供稿：科研事务办公室 编辑：高晨卉 责编：吴淑媛 赵霞