

人工智能数据挖掘在智能车企的规划应用研究

唐明*

(四川农业大学, 四川 成都 611134)

摘要: 随着全球汽车产业在“数据驱动”方面加速变革, 智能车企海量数据资产难以通过传统方式实现价值化利用, AI 数据挖掘成为破局的关键。本文从 AI 数据挖掘技术在智能车企的应用出发, 提出汽车 AI 数据挖掘的关键技术(数据预处理、机器学习、深度学习等)、全价值链规划(研发、设计、生产、运维、服务、用户等)、应用实践(比亚迪全流程数据应用、NVIDIA 安等), 总结 AI 数据挖掘核心技术(数据安全、算法稳定可靠、人才培养等)和应对对策, 通过 AI 数据挖掘规划与应用推动智能车企“数据资产化”, 为汽车产业智能化转型提供借鉴。

关键词: 智能车企; 人工智能; 数据挖掘; 全价值链; 数字化转型

一、引言

1.1 研究背景

目前汽车行业正在经历“电动化、智能化、网联化、共享化”的新一轮变革, 传统的机械车逐步被赋予感知、决策、通信等功能的智能移动互联网所取代^[1]。行业数据指出, 一辆 L4 的智能网联汽车小时的数据量在 500GB 以上, 包括车辆状态数据(电 SOC、电机转速等)、环境数据(激光雷达点云数据、摄像头数据等)、行为数据分析(驾驶加减速习惯、驾驶方向等)^[2]。这些数据是智能汽车公司最核心的资产, 传统的数据处理方式(Excel 表格方式、SQ 方式等)已经无法胜任其“大规模、高维度、快实时”的特点, 导致这些数据无法被利用、无法形成竞争力。

政策与消费者需求驱动 AI 数据挖掘落地。《“十四五”数字经济发展规划》中要求“发展数据驱动的智能网联汽车”; 特斯拉、比亚迪等各大车企都将“数据驱动”写入公司战略, 使用数据挖掘产品与服务; 85%的消费者认为, 应该为其提供定制化的车载体验与安全警示服务^[3]。AI 数据挖掘应用于车企是一个系统工程, 如何系统规划, 是未来智能车企数字化建设的关键点。

1.2 研究目的与意义

1.2.1 研究目的

本文旨在解决 3 个问题, 即: 理清 AI 数据挖掘汽车领域的相关技术, 明确 AI 技术的边界; 建立 AI 技术应用于智能车企全产业链的路线图, 提出可行的路径实施路径; 分析应用过程中的痛点, 提供解决方案, 使车企少走弯路, 降低风险, 提高收益。

1.2.2 研究意义

理论方面, 本文将 AI 数据挖掘技术与汽车产业特点相结合, 填补了汽车产业“技术-场景-价值”衔接研究的空白, 扩展了智能车企数字化转型理论; 实践方面, 案例拆分、框架输出, 供中小车企参考, 减少试错。比如比亚迪的“规划”框架, 缩短研发周期 20%, 降

作者简介: 唐明(1994-), 男, 本科学历, 高级工程师, 研究方向为信息系统工程、软件开发运维等。

通讯作者: 唐明

低售后成本 18%^[4]，具有参考意义。

1.3 国内外研究现状

国外主要围绕技术的实现和商业变现展开研究。特斯拉通过车载传感器收集相关驾驶数据，采用基于深度学习的 FSD 系统，实现从 6 个月的迭代升级为 1 个月的迭代；NVIDIA 聚焦智能驾驶领域的安全痛点，研发全栈式安全产品—halos，从芯片、算法、流程三个层面提升 AI 决策的可靠性。

国内研究以政策牵引和技术领先为主要内容。比亚迪建立“大数据+大算力+大模型”平台，以数据挖掘贯穿研发、生产、运营等流程，在提升产品性能的同时提供更优质的用户体验；研究学者通过文本挖掘分析口碑数据，采用 Stacking 集成学习算法对文本进行情感分类，准确率为 95%、产品改进有明确指导^[5]。

目前有研究已经证实了 AI 数据挖掘的落地性，但仍存在以下问题：缺少全价值链的系统顶层设计，多偏重单一环节应用，缺少数据安全与算法可靠性的解决思路。本文从以下几点展开：缺少全价值链的系统顶层设计，多偏重单一环节应用。

二、智能车企 AI 数据挖掘的核心技术体系

AI 数据挖掘智能车企运用多个技术进行深度挖掘，其架构由底层、算法层、应用层组成，技术层与汽车场景深度融合。

2.1 基础层技术：数据预处理与整合

基础层为应用技术之基，是应对汽车信息“脏、乱、杂”的关键，针对汽车信息源多源多构问题，采用数据清洗技术去除“脏”“乱”“杂”的传感器数据；采用摄像头、雷达、GPS 等多模态融合技术融合多源数据形成单数据视图^[6]；采用主成分分析（Principal Component Analysis, PCA）降维法将高维环境感知数据降维处理，降低计算复杂度；采用时间序列基于滑动窗口的特征提取技术提取车辆工况时序变化趋势。在比亚迪技术的基础上，这一层的可用性从 65%提升到了 92%。

2.2 算法层技术：核心价值挖掘工具

算法层是价值挖掘的核心，不同算法适配不同汽车场景：

机器学习：适合在结构化数据上进行学习，如对零部件运行数据分析，使用随机森林算法预测维修数据。比亚迪基于机器学习建立电池故障预测模型，提前 3 个月预测电池发生衰减的风险，使故障降低 18%；

深度学习：面向非结构化处理数据，例如 CNN 摄像头识别道路交通标识、RNN 驾驶行为时序数据处理识别驾驶危险；

自然语言处理：针对数据信息挖掘，通过 LDA 主题建模和情感分类算法分析用户评论、用户投诉文本，提取产品缺陷信息，为改进提供依据。

2.3 应用层技术：场景落地支撑

应用层将算法结果应用于服务中，包括实时计算和可视化。实时计算技术（例如，Flink）处理车辆实时数据，支持自动驾驶决策、车辆故障预警等场景。可视化技术将挖掘结果应用于运营仪表盘，便于管理层对生产、服务数据的实时监控^[7]。NVIDIA 将芯片级优化技术应用与每秒 300 万亿次安全数据，并将处理延迟保持在 50ms 以下，支持智能驾驶。

三、智能车企 AI 数据挖掘的全价值链规划框架

立足核心技术体系，依托智能车企业业务体系，建立“四环三保”关键环节、落细落

实的整条全产业链部署规划体系。

3.1 核心业务环节规划

3.1.1 研发设计环节：缩短周期，优化性能

运用数据挖掘研发历史数据、用户需求数据改善设计方案、比亚迪建立 AI 设计造型，通过对大量经典案例学习，生产设计概念，根据工程输入，改善设计细节、设计效率、根据底盘仿真工具，建模精准度高，根据实时数据，参数控制，个性化调整、根据用户使用习惯数据、环境数据，建立热舒适性模型，使空调智能感知，自动变化

3.1.2 生产制造环节：提质降本，提升效率

生产过程，利用数据挖掘，实现质量管理、智能化生产。比亚迪利用物联网、大数据改造车间，在线排产、按需生产；利用设备的分析数据、在线检测数据，提升冲压、焊接等工序的连续性、效率等；利用仓储布局的实时监控，提升零部件的摆放、物流调度，降低运输成本。某合资车企利用此种方法，节约一条生产线一年的成本达到 1200 万元^[8]。

3.1.3 运维服务环节：主动预警，优化体验

根据车辆数据及用户使用情况，实现运维服务从响应式到主动创造价值的转变。比亚迪收集近万个整车信号数据，建立健康评估模型，预测空调滤芯及刹车片寿命，实现运维从经验驱动向数据驱动的转变；根据 OTA 技术及数据挖掘的结果，实现整车功能的远程优化，不断地提升用户体验。

3.1.4 用户运营环节：精准服务，提升粘性

挖掘用户数据，精准定制服务。利用驾驶习惯数据，为用户定制专属的驾驶方式，做到千人千面；根据出行路线、出行喜好，为用户提供充电桩、停车场所等场景服务。Stellantis 使车载服务用户数量增长 25%^[9]，用户黏性提升 18%。

3.2 落地保障体系

3.2.1 组织保障：跨部门协同机制

组建“数据挖掘项目组”，包括研发、生产、售后等，消除信息孤岛；设置数据分析师，进行技术支撑、数据培训，建立全员数据文化意识^[10]。

3.2.2 技术保障：统一数据平台

建设“汽车公司数据库中台”，将各个环节数据汇聚共享；采用云计算架构以及高性能芯片（NVIDIA DRIVE AGX），实现海量数据存储，实时计算。

3.2.3 数据保障：规范管理标准

制定数据分类分级标准，划分权利、责任；建立数据质量考核制度，检查完整性、准确性、数据可使用性。

四、典型案例分析

用国内、国外典型应用实例说明 AI 数据挖掘在智能车企中的应用与设计。

4.1 比亚迪：全价值链数据挖掘应用实践

比亚迪“大数据-大算力-大模型”架构下 AI 数据挖掘渗透到全链条：研发端使用 AI 设计工具+仿真模型减少时间 20%；生产端通过智能生产厂改造，提高能效，提高品质；运维端通过健康模型对零部件寿命预测 30%，缩短售后周期。通过客户数据进行能量管理方案调优，平衡动力与能耗，提升客户体验。这是全价值链架构的落地。

4.2 NVIDIA：算法安全数据挖掘解决方案

针对智能驾驶 AI 算法的可靠性痛点, NVIDIA 推出 Halos 系统: 算法层开发“安全熵值”评估体系, 用生成式 AI 模拟 2.8 亿种攻击场景; 流程层建立 11 道安全闸口, 实现从数据标注到 OTA 的全链条防护。通用汽车应用后, SuperCruise 系统误触发率直降 67%; Gatik 物流车夜间物体识别准确率从 82% 飙升至 99.3%, 为算法可靠性保障提供了有效路径。

五、应用挑战与解决方案

5.1 核心挑战

5.1.1 数据安全风险突出

智能驾驶汽车数据中包含用户位置、驾驶行为等个人信息, 数据传输与存储存在风险; 另外, 数据受到污染也会导致结果分析不准确, 影响决策制定。某智能驾驶汽车企业曾因数据安全问题的暴露了 10 万用户驾驶数据, 造成企业形象不佳。

5.1.2 算法可靠性不足

现有的算法都是根据特定场景进行的, 无标线等极端天气场景的准确率降低, 容易造成自动驾驶决策失误; AI 的“黑箱”效应, 故障溯源困难, 易引发安全隐患。

5.1.3 复合型人才短缺

行业需要的是 AI 技能、汽车专业技能和数据技能的复合人才, 高校相关专业培养的落后, 企业内部人员转化难度大, 不利于技术的实施。

5.2 解决方案

5.2.1 构建全栈式数据安全体系

硬件层面使用 PUF (Physical Unclable Function) 实现芯片的唯一“数字指纹”; 软件层面使用区块链升级 OTA 存证, 实现数据加密; 流程层面实现数据“消毒”, 屏蔽恶意数据。严格遵守《数据安全法》, 设定数据采集范围, 保护隐私。

5.2.2 提升算法鲁棒性与透明度

通过虚拟场景孪生技术, 虚拟极端事件, 扩大训练数据集; 通过“加强安全学习”技术, 使系统面对未知威胁时进入安全模式; 通过决策过程跟踪系统, 减少“黑箱”效应。

5.2.3 完善人才培养体系

企业与高校合作特色专业, 开设“汽车 AI 数据挖掘”; 内部进行交叉培训, 选派精英人才进行培养; 企业行业联盟资源共享。

六、结论

本文对 AI 数据挖掘在智能车企中的设计与实现进行研究, 得出以下结论: 首先, 构建汽车 AI 数据挖掘技术系统的基础层、算法层、应用层, 各层次技术支持价值挖掘; 其次, “四环节+三保障”的整体价值链设计框架可以更好地支持在研发、设计、生产、运行维护环节的人工智能 AI 数据挖掘; 再次, 安全问题、算法问题、人才问题为主要痛点, 全栈式安全、算法优化、人才培养可以很好的解决问题。

AI 数据挖掘是智能车企实现数字化转型的重要手段, 探索 AI 数据挖掘融合汽车数据应用和联邦学习数据共享应用, 为高质量发展赋能。

参考文献:

- [1] 工信部. 智能网联汽车产业发展规划 (2021-2025 年) [Z]. 2020.
- [2] 中国汽车工程学会. 2024 年中国智能网联汽车数据白皮书 [R]. 北京: 中国汽车工程学会, 2024.
- [3] 国家互联网信息办公室. 汽车数据安全管理办法 (试行) [Z]. 2021.

- [4] 廉玉波. IoE 时代智能电动汽车数据智能应用探索[J]. 电子工程专辑, 2025 (6) :45-51.
- [5] 张明, 李丽. 基于文本分类的新能源汽车满意度研究——以比亚迪与特斯拉为例[J]. 数据分析与决策, 2025 (5) :78-85.
- [6] 王建华, 刘洋. 智能网联汽车多源数据融合技术研究[J]. 汽车工程, 2024, 46 (3) :412-418.
- [7] 李敏, 陈浩. 构建与实施车企数据中台[J]. 数字技术与应用, 2024 (2) :135-137.
- [8] 人工智能应用于汽车制造生产过程的人工智能应用研究[J]. 制造业自动, 2024, 46 (4) :89-93.
- [9] Stellantis 集团. 2024 年用户运营数据报告[R]. 荷兰:Stellantis 集团, 2024.
- [10] 刘芳. 智能车企数字化转型挑战与对策[J]. 企业经济, 2024 (3) :76-82.

Research on the planning and application of artificial intelligence data mining in intelligent automobile enterprises

TANG Ming*

(Sichuan Agricultural University, Chengdu, Sichuan 611134, China)

Abstract: As the global automotive industry accelerates its "data-driven" transformation, intelligent automakers face challenges in effectively leveraging massive data assets through traditional methods. AI data mining has emerged as a critical breakthrough. This paper explores the application of AI data mining technologies in smart automotive enterprises, proposing key technologies (data preprocessing, machine learning, deep learning), comprehensive value chain planning (R&D, design, production, operations, services, user engagement), and practical implementations (BYD's end-to-end data applications, NVIDIA security solutions). It summarizes core AI data mining technologies (data security, algorithm stability, talent development) and corresponding strategies. Through strategic planning and practical applications of AI data mining, this study drives the "data assetization" of intelligent automakers, providing valuable insights for the industry's intelligent transformation.

Keywords: Smart car manufacturers; Artificial intelligence; Data mining; Full value chain; Digital transformation