

术弹性分配资源, 高峰时段(如早晚通勤地铁内): 自动激活基站冗余载波, 将单基站容量提升 2-3 倍, 同时将相邻基站的覆盖范围微调(如缩小边缘小区覆盖), 避免信号干扰, 保障用户体验; 此时虽能耗略有上升, 但单位业务量能耗(每 GB 流量能耗)下降。低谷时段(如深夜居民区): 关闭非核心载波, 将相邻基站的覆盖范围适度扩大(如合并小型小区), 减少活跃基站数量, 同时降低基站发射功率, 实现“少设备高覆盖”, 能耗降低 20%-30%。

数据中心算力调度: 基于“负载预测”的算力池化管理。利用 AI 预测未来 24 小时数据处理需求(如电商大促前算力需求激增), 提前将闲置服务器的算力整合到“算力池”, 分配给高需求业务; 非高峰时段则将部分服务器转入“低功耗休眠模式”(仅保留核心存储功能), PUE(能源使用效率)可从 1.5 降至 1.2 以下(PUE 越接近 1, 能耗越优)。将高频访问数据(如短视频、本地生活服务数据)存储在靠近用户的边缘数据中心, 减少跨区域长途传输的能耗(长途传输每公里能耗约为本地传输的 5 倍), 同时降低网络延迟。

用户侧服务与引导: 从“终端到行为”的全链条低碳干预。用户侧是移动通信全链路的“末端环节”, 其终端设备能耗、用网行为直接影响整体碳足迹, 优化措施聚焦“技术降耗+行为引导”双管齐下:

低功耗终端推广与技术优化: 从“硬件到传输”降低终端能耗。联合终端厂商推出“通信节能认证终端”: 要求手机、IoT 设备(如智能电表、摄像头)采用低功耗芯片(如高通骁龙 8 Gen3 节能版)、优化射频模块设计, 待机功耗降低 15%-20%; 同时推动终端支持“绿色通信协议”(如 5G NR 节能模式), 在信号弱区域自动降低搜网频率, 减少耗电。

优化内容传输技术: 对视频、直播等大流量业务采用“动态码率编码”, 根据用户网络质量自动调整画质(如 4G 网络下输出 720P 画质, 5G 网络下输出 4K 画质), 避免“高码率低画质”的流量浪费; 同时推广“边缘缓存”技术, 将热门视频提前存储在用户附近的基站, 减少重复传输能耗, 用户端流量能耗可降低 10%-15%。

4.4 用户用网行为引导: 通过“激励+教育”培养低碳习惯

运营商 APP 新增“低碳用网积分体系”: 用户开启手机“夜间省电模式”、使用 WiFi 替代蜂窝网络、参与旧终端回收等行为, 可兑换流量或话费优惠, 引导用户主动降低终端能耗。

开展“低碳用网公益宣传”: 通过短信、APP 弹窗推送节能小贴士(如“关闭后台冗余 APP 可降低手机耗电 30%”“视频自动播放清晰度设为‘自动’更节能”), 提升用户低碳意识, 形成“企业引导+用户参与”的减排合力。

5 供应链与生态协同—贯穿“全生命周期”的低碳生态构建

移动通信设备的“生产-采购-使用-报废”全生命周

期占行业总碳排放量的 40% 以上(含上游原材料生产、下游回收处理), 优化措施聚焦“产业链协同”, 推动全流程低碳化:

5.1 低碳供应商筛选与管理

从源头控制设备碳足迹, 建立“通信设备供应商碳足迹评估体系”: 将设备生产过程中的能耗(如服务器芯片制造能耗)、原材料碳排放(如光纤生产的塑料原料碳排放)纳入采购考核, 优先选择碳足迹低于行业均值 10% 的供应商; 对核心设备(如基站、服务器)要求提供“碳足迹报告”, 实现全链条可追溯。

推行“集中采购+长期合作”模式, 通过大规模集中采购降低供应商生产环节的冗余能耗(如同一型号基站批量生产比小批量生产能耗降低 15%), 同时与供应商签订“低碳合作协议”, 要求其 3 年内将生产环节碳排放降低 20%, 否则取消合作资格。

5.2 设备全生命周期管理与跨行业协同

减少报废环节污染与能耗, 建立“通信设备回收拆解体系”: 对退役基站、服务器、手机等设备, 通过专业机构进行分类拆解, 可回收金属(如基站外壳的铝合金)、芯片、电池等部件, 回收率达 80% 以上; 对不可回收部分采用环保处理技术(如无害化焚烧、降解), 避免电子废弃物污染, 同时减少新设备生产的原材料需求(每回收 1 吨铝合金可节约 9 吨铝土矿, 减少 5 吨碳排放)。

跨行业合作共建“绿色能源-通信一体化平台”, 与光伏、风电企业合作, 在基站、数据中心屋顶及周边区域建设分布式光伏电站, 优先满足通信设备用电需求(如某地区基站光伏供电占比达 30%, 年减排 1200 吨二氧化碳); 同时与电网企业合作, 将数据中心、基站纳入“虚拟电厂”体系, 在电网用电高峰时段(如夏季午后)降低通信设备非核心负载, 助力电网削峰填谷, 获取绿电补贴, 实现“节能+收益”双赢。

6 结语

只有当基础设施的“绿色硬件”与运营端的“智能管理”形成协同可再生能源供电配合负荷预测运维——才能实现“1+1>2”的减排效果, 推动移动通信全链路从“被动节能”向“主动降碳”转型, 最终为“零碳通信生态”筑牢核心支撑。并且通信技术的飞速发展, 6G 时代的到来已不再遥远, 6G 网络不仅要满足人们对高速、低延迟和海量连接的需求, 通信网络的低碳化转型显得尤为迫切。

参考文献

- [1] 移动通信工程中的光纤传输系统设计与优化[J]. 查双全. 信息与电脑(理论版), 2024(17)
- [2] 移动通信信号安全传输的保障策略分析[J]. 张超. 集成电路应用, 2024(08)
- [3] 通信网的光纤传输安全和保护探析[J]. 张兴之. 中国新通信, 2023(13)

The application of big data technology in enterprise artificial intelligence

Jianjun Wu

Shanxi Longxing Network Technology Co., Ltd., Lvliang, Shanxi, 044201, China

Abstract

In the wave of digital transformation, enterprises' AI development heavily relies on massive and diverse data support. As a core tool for data governance and value extraction, big data technology has become a critical infrastructure for implementing enterprise AI. Taking the "51 Good Electric AI Platform Development and Promotion" project as a case study, this paper systematically analyzes specific application scenarios of big data technology in enterprise AI, covering data collection and preprocessing, algorithm model optimization, adaptive learning enhancement, and multi-modal interaction implementation. The research demonstrates that big data technology addresses core pain points in enterprise AI such as data quality, computational power adaptation, and real-time responsiveness, facilitating the transition of AI models from "theoretically feasible" to "operationally viable." This significantly improves operational efficiency and service intelligence levels. Additionally, by integrating novelty search conclusions with industry practices, the paper identifies future trends in big data-AI integration, providing technical references for enterprise digital transformation.

Keywords

big data technology; enterprise artificial intelligence; data governance; algorithm optimization; adaptive learning

大数据技术在企业人工智能中的应用

武建军

山西龙行网络科技有限责任公司, 中国·山西 吕梁 033000

摘要

在数字化转型浪潮下,企业人工智能(AI)的海量多元数据支撑型发展高度依赖数据,而作为数据治理与价值挖掘核心手段的大数据技术,已成为企业AI落地的关键基础设施。本文以“51好电AI平台研发与推广”项目为实践案例,结合国内相关技术研究与企业应用现状,系统分析大数据技术在企业AI中含数据采集与预处理、算法模型优化、自适应学习能力提升及多模态交互实现等维度的具体应用场景。研究表明,大数据技术通过解决企业AI的数据质量、算力适配、实时响应等核心痛点,助力AI模型从“理论可行”走向“业务可用”,显著提升企业运营效率与服务智能化水平。同时,结合查新结论与行业实践,指出大数据与AI融合的未来趋势,为企业数字化转型提供技术参考。

关键词

大数据技术;企业人工智能;数据治理;算法优化;自适应学习

1 引言

随着人工智能技术产业化落地阶段的到来,企业对AI的需求已从“技术探索”转向“价值创造”。但AI模型训练等海量高质量数据需求场景下,约80%数据质量与处理能力不足的企业面临AI项目落地难题。在此背景下,具备数据采集等核心能力的大数据技术,已成为连接企业业务数据与AI价值的桥梁。以“数据驱动决策”为核心目标的企业AI需经历“数据输入—模型训练—迭代优化—业务输出”闭环,其中直接影响AI模型性能与业务适配度的数据输入

等环节至关重要^[1]。以“51好电AI平台”为例,其依赖大数据技术实现多源异构数据治理与算法模型优化的多元服务已形成实践案例。目前,京东等国内企业虽已布局“大数据+AI”融合应用,但多聚焦单一环节,缺乏全流程一体化技术架构。本文基于“51好电AI平台”资料与国内研究成果,剖析大数据技术在企业AI中的应用逻辑与实践路径,为企业构建高效可持续AI体系提供参考。

2 大数据技术为企业人工智能提供的核心支撑

2.1 数据治理

AI模型的训练与优化依赖高质量数据,而企业数据往往存在“碎片化”“非结构化”等问题。以电商平台为例,其包含浏览记录类结构化数据、客服对话类文本非结构化数

【作者简介】武建军(1978-),男,中国山西吕梁人,本科,高级工程师,从事互联网,人工智能研究。

据、商品图片类图像非结构化数据的用户数据，若未经治理直接输入模型，将导致模型精度下降。大数据技术通过“采集—清洗—标准化—存储”的全流程治理，为 AI 提供“干净、可用”的数据。以 51 好电 AI 平台为例，其采用 Flume 与 Kafka 等大数据分布式采集技术，实时抓取电商场景下用户点击、收藏、购买等行为数据、用户咨询录音类语音交互数据、商品评价与客服对话类文本数据，并通过缺失值填充与异常值剔除等数据清洗算法处理噪声数据；同时，利用 SparkMLlib 等数据标准化工具将非结构化数据转化为语音数据转写文本向量、图像数据提取特征矩阵等结构化格式，最终存储于 HadoopHDFS 等分布式数据库中，为后续 AI 模型训练提供统一数据源^[2]。此外，大数据技术还通过“数据集治理”提升数据复用性。

2.2 算力调度

AI 模型尤其是深度学习模型的训练需消耗大量算力，企业爆发式增长的数据量进一步加剧算力需求。大数据技术通过分布式计算框架与算力调度算法实现算力资源的高效分配，缩短模型训练周期。从技术实践来看，51 好电 AI 平台基于大数据技术构建“算力调度中心”，采用 Kubernetes 容器编排引擎与 YARN 资源管理系统将 CPU、GPU 等算力资源虚拟化，根据模型训练需求动态分配资源，在语音识别模型训练时调度 GPU 资源加速深度学习计算，在批量数据统计处理时分分配 CPU 资源执行并行计算。这种调度模式不仅降低算力浪费使资源利用率提升 30% 以上，还将模型训练周期从传统的“数天”缩短至“数小时”。国内相关研究也印证这一逻辑：焦点科技在其“基于电商应用的人工智能平台”中通过大数据算力中心实现对推荐模型的高效训练；京东 AI 平台依托云计算与大数据融合架构支撑零售、物流等多场景的 AI 模型训练，其算力调度效率较传统架构提升 50%。

2.3 实时分析

企业 AI 业务应用需具备支撑实时推荐、智能客服等场景的“低延迟”特性。电商平台实时推荐系统需在 1 秒内依据用户实时行为数据生成避免转化机会流失的推荐结果。大数据技术借助 Flink、Spark Streaming 等流处理框架快速分析高并发实时数据为 AI 模型提供“即时决策”能力。51 好电 AI 平台的“智能聊天系统”依靠大数据实时分析技术在用户咨询时通过流处理框架实时抓取输入文本、调用 NLP 模型解析意图并关联历史数据生成个性化回复且将响应延迟控制在 500 毫秒以内。其“商业分析模块”通过实时分析用户行为数据动态调整商品推荐策略使商品点击率提升超 20%。深圳追一科技在“基于 NLP 的电商客服平台”采用类似技术实时处理用户咨询数据使智能客服问题解决率提升至 85% 并减少人工客服 30% 工作量进而证明实时分析技术能提升企业 AI 业务价值。

3 大数据技术在企业人工智能中的具体应用场景

3.1 算法模型优化

算法模型是企业 AI 的“大脑”，大数据技术通过数据增强、参数调优、模型集成三大手段优化适配企业业务需求的模型精度与泛化能力。从数据增强来看，企业 AI 常面临“数据量不足”问题，51 好电 AI 平台“云绘画模块”仅依赖少量用户创作数据训练会导致模型生成单一图像风格，此时大数据技术通过图像翻转、文本同义替换等数据扩充算法生成虚拟训练数据，结合真实数据共同训练模型以提升 25% 模型泛化能力。在参数调优方面，深度学习模型学习率、卷积核大小等 AI 模型参数直接影响性能，传统人工调优方式效率低且精度差，大数据技术通过 Hyperopt、Optuna 等自动化调优工具基于历史训练数据构建参数优化模型并实时迭代参数组合，51 好电 AI 平台采用该技术将模型参数调优时间从“数天”缩短至“数小时”且同步提升 15% 模型准确率。模型集成通过大数据技术整合多个单一模型优势以提升整体性能，51 好电 AI 平台“商业分析模块”将销量预测用线性回归模型、用户画像用随机森林模型、趋势分析用深度学习模型通过 Stacking 等大数据集成算法融合，将销量预测误差降低至 8% 以下且较单一模型提升 30% 性能。国内相关研究也验证这一应用效果，网络文献《AI 深度学习模型优化算法》指出大数据技术应用梯度下降法、批标准化、正则化等算法可提升 40% 深度学习模型训练效率与 20% 泛化能力；焦点科技电商 AI 平台通过协同过滤+深度学习集成推荐算法提升 25% 商品推荐准确率^[3]。

3.2 自适应学习

企业具有随季节、促销活动变化的电商平台用户偏好等动态业务需求与用户行为，AI 模型若无法实时调整将导致业务适配度下降。大数据技术借助实时数据反馈与增量学习算法赋予 AI 模型实现持续进化的自适应学习能力。51 好电 AI 平台的自适应学习实践具有代表性，其先通过大数据实时采集对推荐商品的点击/跳过、对智能客服回复的满意度评分等用户反馈数据并将其转化为模型优化的监督信号；再采用基于 Spark 的增量训练框架等增量学习算法，在不重新训练整个模型且无需中断服务的前提下将新数据融入现有模型以实现轻量化迭代，如用户对“生活娱乐 AI”模块内容推荐满意度下降时模型可通过增量学习快速调整推荐策略；最后引入 MAML 算法等元学习技术，使模型能基于历史学习经验在少量数据下快速上线以适配新增“角色助手”服务等新业务场景。这种自适应学习能力显著提升 51 好电 AI 平台的业务适配度，据查新资料显示该平台在用户行为变化时模型调整响应时间从 24 小时缩短至 1 小时且用户满意度维持在 90% 以上。京东 AI 平台也采用类似技术，其零售场景推荐模型通过实时分析用户行为数据动态调整