运营商AI能力建设及演进探讨

Discussion on Construction and Evolution of Operators' AI Capability

陈俊明[1,2]，张　洁[3]，左　罗[1,2]（1. 南京中兴新软件有限责任公司，江苏 南京 210012；2. 移动网络和移动多媒体技术国家重点实验室，广东 深圳 518055；3. 南京师范大学中北学院，江苏 镇江 212300）

Chen Junming[1,2]，Zhang Jie[3]，Zuo Luo[1,2]（1. Nanjing ZTE New Software Co.Ltd.，Nanjing 210012，China；2. State Key Laboratory of Mobile Network and Mobile Multimedia Technology，Shenzhen 518055，China；3. Nanjing Normal University Zhongbei College，Zhenjiang 212300，China）

摘　要：

5G时代运营商面临各种挑战，亟需引入AI能力进行应对。对于AI能力的建设，提出了以AI中台为核心的整体架构，给出了AI中台的建设模式、模型重训练、模块互联互通等方面的建议，并针对AI安全、隐私保护及数据治理进行了探讨。最后，对AI能力的演进给出了中短期发展建议。

关键词：

AI；中台；模型重训练；互联互通；安全；数据治理

doi：10.12045/j.issn.1007-3043.2021.12.017

文章编号：1007-3043（2021）12-0083-06

中图分类号：TN919

文献标识码：A

开放科学（资源服务）标识码（OSID）：

Abstract：

In 5G era，operators are facing various challenges，and the introduction of AI is necessary. For the construction of AI capability，The overall intelligent architecture with AI middle platform as the core part is proposed，and suggestions on construction mode of AI middle platform，AI model retraining，module interconnection and interworking are given，and AI security，privacy protection and data governance are discussed. Finally，suggestions on the evolution of AI capability in the near future is put forward.

Keywords：

AI；Middle platform；Model retraining；Interworking；Security；Data governance

引用格式：陈俊明，张洁，左罗. 运营商AI能力建设及演进探讨[J]. 邮电设计技术，2021（12）：83- 88.

0 引言

近年来以5G、AI、云计算为代表的新技术迅猛发展，运营商逐步从主要服务于人转向全面服务于整个社会。人与人通信的单一模式逐渐演化为人与人、人与物、物与物的全场景通信模式，业务场景更加复杂。业务场景的复杂性将带来对SLA的差异化需求以及与之配套的网络管理的复杂性。2B方面，5G需应用于自动驾驶、工业控制、水表电表的自动抄表、智慧园区、智慧医疗、智能交通、智慧教育等；2C方面，5G需应用于云游戏、AR/VR等。

要支撑这些新业务，运营商面临如下的挑战。

a）新业务开发速度的挑战：传统方式下，由基础设施直接提供业务，相关能力竖井状散落在各个具体的业务中，新业务开发周期长。

b）云网拉通的挑战：OTT通过公有云、私有云、混合云、异构云，为各行业客户提供多环境、多形态、随需部署的多样化云服务，给电信运营商的运营带来极大的竞争压力；同时很多业务的提供还需要将云和网打通，但目前云网协同尚在推进中，业务的交付周期

收稿日期：2021-11-05

长、业务质量保障方面仍存在不足。拉通IT、CT、DT、OT能力,提供一体化服务,是电信业发展的必然。

c)运维的挑战:运营商的网络很长一段时期内都会是多制式(2G/3G、4G、5G)共存的环境,由此带来了协同和互操作难度,同时网络分层解耦架构带来故障定界定位困难,虚拟化/云化网络的动态变化带来资源统一调度和管理挑战等。

AI在特性挖掘、深度数据分析、策略动态生成等方面具备很大优势,将AI技术引入通信网络可以助力电信运营商构筑更加灵活、高效的信息基础设施,从而进行业务流量预测、设备的预防性维护和资源优化分配,减少重复性人工操作,可以更快速地拉通云网业务,提升新业务的开发速度。目前,运营商已在AI领域积极地开展实践。

1 AI能力构建

1.1 云网融合智能化整体架构

云网智能化可以基于基础设施层、管控层、跨域三层网络架构实现,可以将AI能力模块化设计,随需植入云网基础设施层、单域管控层和跨域运营层。顶层架构如图1所示。

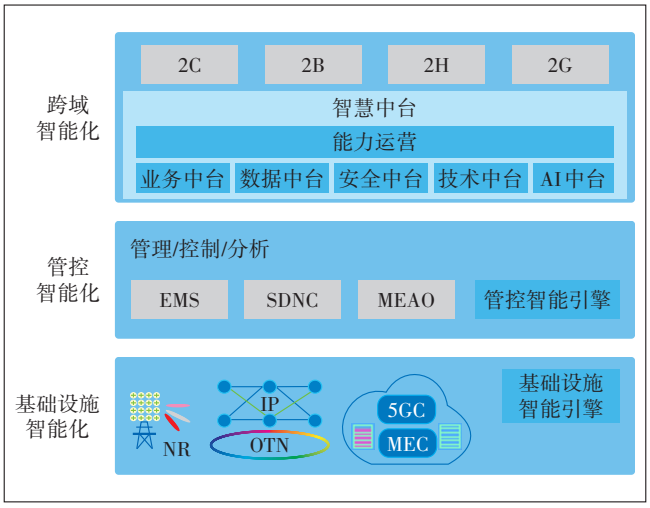

图1 运营商智能化整体架构

运营商的AI能力打造首先体现在中台打造上,AI中台承担着AI能力"大脑"的作用,包括AI模型的集中训练、全局推理和AI模型市场等功能。AI中台管理着云网各层面的AI模型,通过与管控智能引擎、基础设施智能引擎的协作实现云网智能闭环。在智慧中台内部,AI中台将AI能力提供给能力运营中心、业务中台、数据中台、技术中台及安全中台。在能力运营中心,可以根据用户的喜好、调用行为给用户推荐适合调用的能力;在业务中台,可将AI能力用到具体的业务中,进行云网的端到端运维,CDN中热点视频的边缘推送,产品的质检、水质的监测、园区的安防等;在数据中台中,可以使用AI能力进行数据质量的检查,对异常数据进行识别,对缺失数据进行补充回填;在技术中台中,可以使用AI能力进行资源的调度,减少资源消耗;在安全中心中,可以使用AI能力进行恶意软件的检测,识别攻击流量。AI中台需处理的数据量大,对算力要求高,对实时性要求相对低,需要集群部署。

运营商的AI能力随需嵌入管控层,形成管控智能引擎,可以快速与现有的运维管控系统相结合,增强云网单域的管、控、析能力,实现单域的智能化,可应用于云网单域告警分析、基站智能节能等场景。这些场景需处理的数据量中等,对算力的要求也适中,实时性要求相对较高,可以使用少量服务器进行部署。

运营商的AI能力也可以嵌入基础设施层,形成基础设施智能引擎,可植入云网基础设施(如基站)实现高实时智能策略,适用于无线动态频谱分配、5GC电信云动态扩容等场景,这些场景需处理的数据量相对较小,对处理的实时性要求最高,可将AI能力嵌入基础设施进行部署。

1.2 AI中台模型开发工具

AI中台模型开发工具需从易用性角度出发,支持基于AI开发工具低码或无码开发,沉淀多样可视化算子,通过简单的拖拽完成从数据挖掘到模型生成的过程,通过比较不同模型ROC曲线、F1值等选择最优模型,利用交互式操作减少工具的使用难度,提升模型的开发效率。

在数据预处理方面,通过散点图、折线图、相关系数热力图、分类聚类雷达图等方便快速发现数据规律,从而为特征工程、模型选择提供帮助。在训练过程中通过模型损失值的变化实时显示、实时中断回滚、自动故障恢复及时调测程序,缩短模型的训练时间。

通过打造简单易用的工具,让更多的业务人员能够利用AI工具来解决业务问题,从而降低AI的使用门槛。

1.3 AI中台建设模式

在AI中台建设中,集团公司负责AI中台集中建设、集约化建设AI能力,构建整个集团公司内的模型

市场。

省分的AI中台建设分为2种情况，一种是对于AI使用需求较少和没有实时AI使用场景的省分，可以分权分域地使用公司的统一AI平台中的部分资源；另一种是公司在省分建设拉远AI中台。不论是哪种方式，省分都可以使用公司发布的模型在生产系统中进行应用，省分也可按需迭代优化模型或者发展省分特色模型并贡献给公司。集团公司、省分在AI中台的分工协同如图2所示。

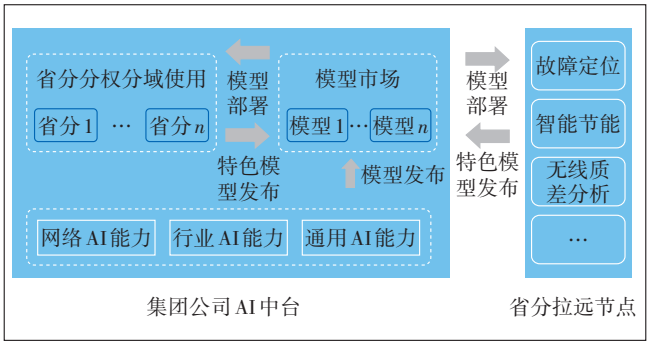

图2　集团公司、省分AI中台分工协同

1.4 AI能力贯通与模型重训练

AI能力贯通主要是通过将AI模型部署到不同层级来实现。对于模型部署，AI中台训练完成的模型按需下发至网络各层，被不同层的系统集成使用。集团AI中台可以向省分AI中台按需下发通用业务模型，省分AI中台可以下发本地特色模型至管控/基础设施智能引擎。

为了确保使用AI后网络的质量还在合理范围之内，需要对AI模型执行的结果设定正确率阈值，在模型推理正确率不能达到要求时，需要有非AI的方案作为备用方案。

AI模型运行一段时间后推理正确率可能不能满足要求，这其中可能有多种原因，如使用者行为的变化、业务配置的变化、数据的变化、业务软件版本的变化、基础设施的变化等，这些情况都需要进行模型重训练。模型重训练分为在线训练及离线训练2种情况，在线训练使用实时流数据进行训练，适用于数据特征快速变化的场合，对算力资源的需求相对高；离线训练使用非实时数据进行训练，适用于数据特征稳定的场合，对算力资源的需求相对低。离线训练也需要定期进行重训练以保证模型的正确率，在系统能够监控模型应用正确率时，还可以设定模型应用的正确率阈值，当正确率低于某个阈值时触发模型的重训练。当然，对模型应用正确率的监控同样适用于在线训练，在当其正确率低于某个阈值时需要重新提取特征/选择其他AI算法或回退到非AI处理方式。AI中台需要考虑支持离线训练和在线训练2种方式，具体场景，初期以离线训练为主，逐步过渡到在线训练方式。智能化能力贯通与重训练的结构如图3所示。

1.5 合作共建AI算法体系

AI可以用于运营商的云和网，可以用于赋能行业算法等；行业的算法比如自动驾驶算法、水质监控算法、水泥的下料口堵塞检测算法、钢铁的淬火温控算法等等，这些能力只凭运营商一己之力无法完全实现，需要与合作伙伴共建。AI算法体系如图4所示。

运营商在与合作伙伴共同建设AI能力过程中，不同模块/系统的互联互通不可避免，需要涉及到数据存

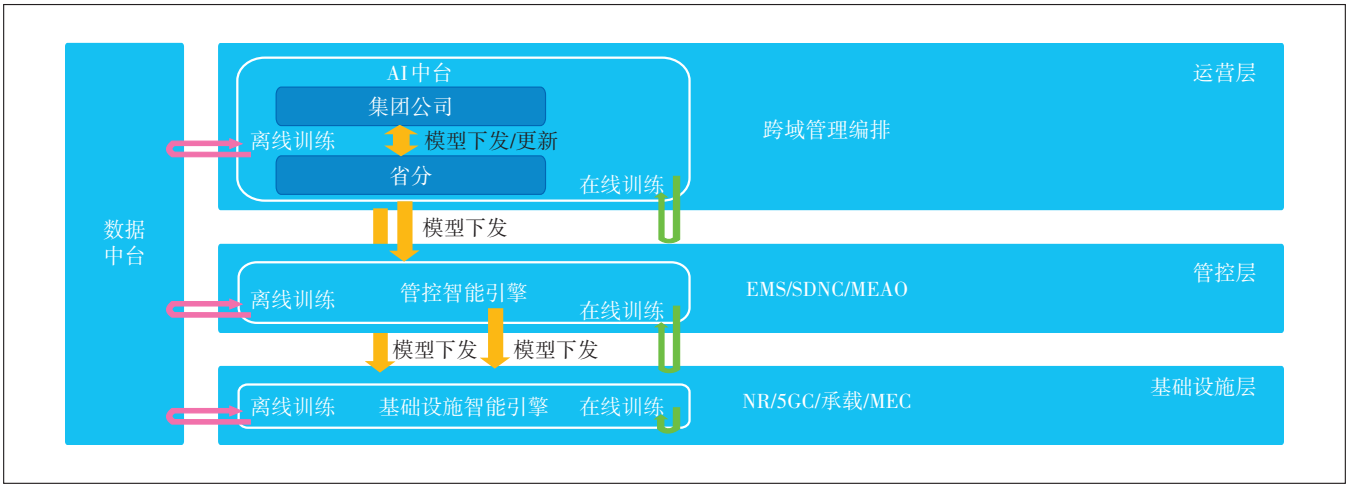

图3　智能化能力贯通与重训练的结构

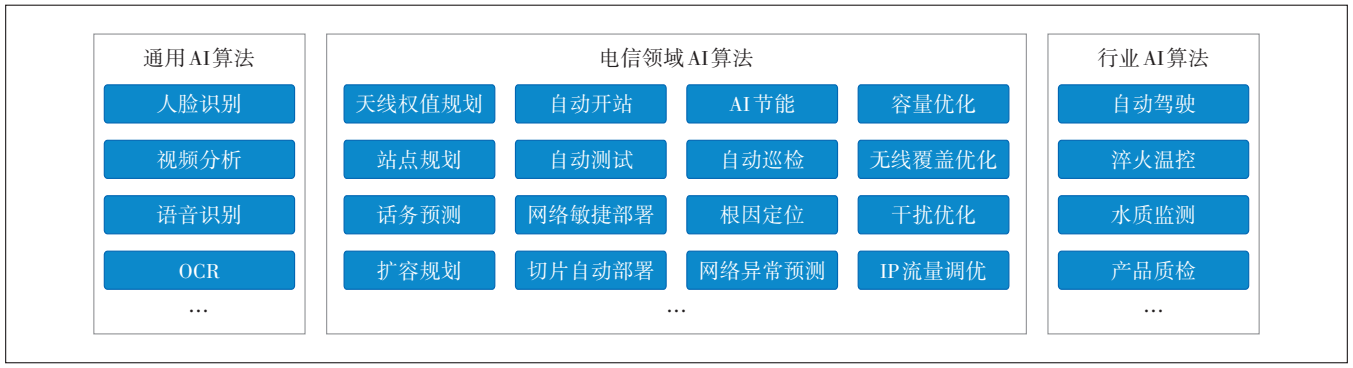

图4 合作构建算法体系

储接口、离线数据访问接口、在线数据访问接口、模型训练调用接口、模型发布接口、模型部署接口、模型访问接口和智能命令接口。运营商可针对这些接口制定规范使得互联互通有章可循,提升对接效率。图5给出了互联互通接口示意。

1.6 AI能力建设配套举措

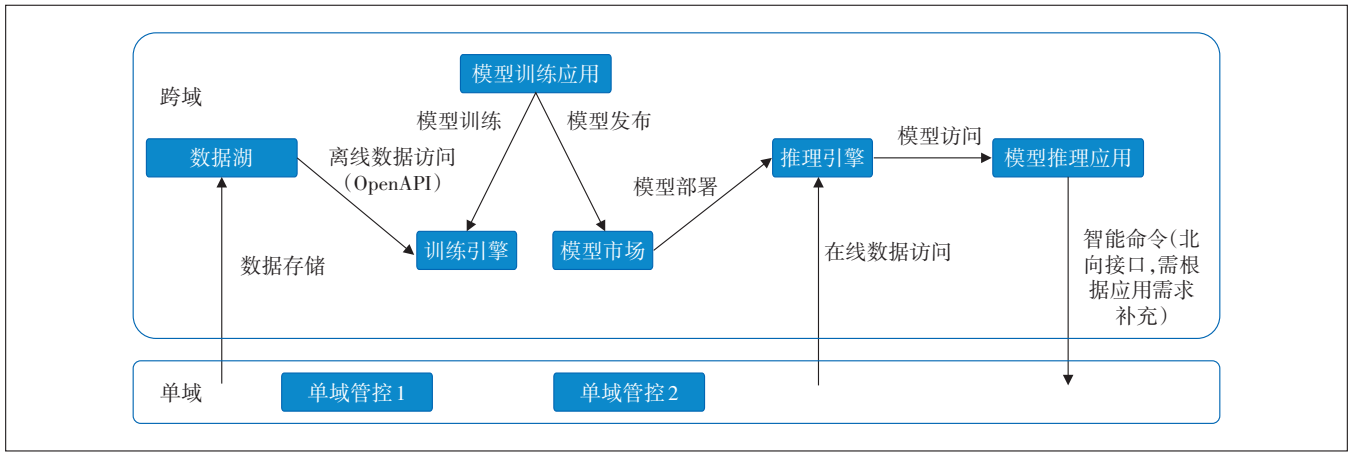

图5 互联互通接口

1.6.1 组织方面

需要汇集AI和各领域业务专家持续打造公司AI能力中心,负责公司统一的AI平台建设、场景能力规划、AI模型的研发、AI产品的开发和模型市场建设。AI能力中心还需负责制定能力开放标准,接口规范,以及协同省分与公司AI能力。省分层面需要展开AI应用创新试点,推动AI模型成熟并复制推广。图6给出了AI团队组建示意。

除了集团公司和省分两级AI团队外,还需要将AI人员嵌入到业务开发团队中,便于消除AI人员与业务人员之间的隔阂,把握业务真实需求,采用最合适的AI算法来构建模型,以及对AI模型的效果进行准确评判。

1.6.2 人才方面

需要储备AI算法专家、大数据专家。此外,AI/大

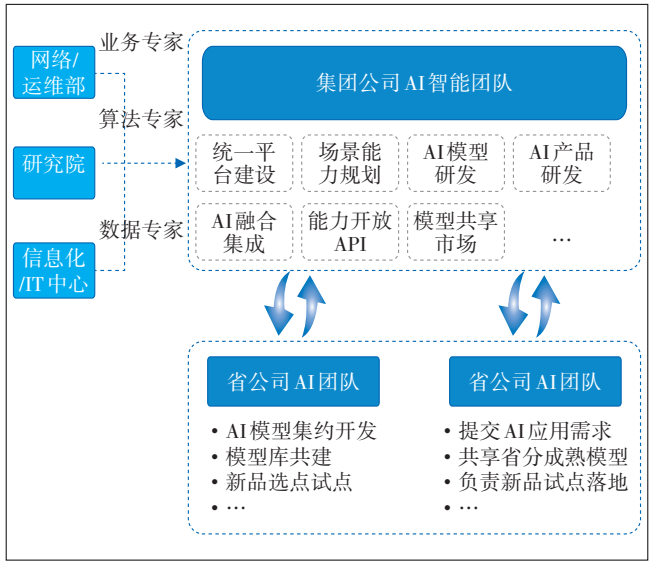

图6 AI团队组建

数据专家需要具备一定的业务能力，业务专家需要具备一定的AI/大数据能力便于协同开展工作，相关能力可通过内外部培训和实战持续改进。

1.6.3 考核方面

需要在考核中对人员工作内容进行调整，人员主要工作要适配AI和自动化的要求，要将经验固化为脚本，要搜集数据训练/重训练AI模型，在某个AI场景的应用初期要对AI处理的结果进行人工抽查并修正。此外还要探索制定对应的新业务维护和运营流程。

1.7 AI能力构建的挑战

AI能力构建存在一些挑战，这跟AI技术本身的发展现状息息相关，主要集中在数据标注效率不高、安全隐私保护挑战和模型可解释性等方面。

1.7.1 数据标注

AI模型训练中有高达70%以上精力花在数据准备和处理上，数据质量差、数据打标效率低是主要问题。应对的办法是组建标注团队，进行标注众筹，并采用机器自动标注和人工检查结合的方式，逐渐提高自动标注水平。同时，充分利用现有带标数据，如故障工单系统数据等。

1.7.2 安全性

AI系统可能遭受各种攻击，如闪避攻击（在正常样本上加入人眼难以察觉的微小扰动，以使AI模型出错）、药饵攻击（污染训练数据，使AI模式出错）、后门攻击（篡改模型，加上了后门）和模型窃取攻击（多次调用AI推理识别接口以窃取AI模型）。对于闪避攻击，需要增强模型本身的健壮性；对于药饵攻击，需要控制对训练数据的采集、过滤数据、定期对模型进行重训练甚至使用实时数据在线训练等一系列方法；对于后门攻击，需要对AI模型做适当的变换；对于模型窃取，可以对训练数据加密、加噪和模型加噪。总的来说，运营商网络运维和2B服务场景隔离性相对更高、被攻击的可能性相对小，对公众运营的2C场景受攻击的可能性相对大。闪避攻击、药饵攻击、后门攻击都会影响AI模型的准确性，对药饵攻击、后门攻击可以通过安全措施的加强来减缓甚至归避，而闪避攻击需要学界不断地研究促进AI算法本身的进步。

1.7.3 隐私和数据治理

AI模型训练过程中会涉及到大量的数据，容易造成用户的隐私泄露，而不准确的数据可能造成偏见。为防止用户的隐私泄露，需要遵守有关法规要求，如《个人信息保护法（草案）》、欧盟GDPR等，进行数据脱敏（加密、匿名化、差分隐私）、分级分类授权使用；需要构建体系化安全系统，记录数据处理的全流程，加强数据访问协议的管理，严格控制数据访问和流动的条件，确保收集到的信息不被非法利用。对于数据不能出本地的情况，可引入联邦学习，在不占有数据的基础上训练出AI模型。对于AI系统可能造成的歧视弱势群体的情况，需剔除数据中错误、不准确和有偏见的成分。

1.7.4 模型可解释性

有些模型是通过算法直接从数据中创建，人们无法理解如何将变量组合在一起进行预测。模型可解释受关注的地方主要在用户体验方面，比如信息流推荐、商品推荐等。目前主要做法是将不可解释的模型用可解释的模型如决策树等替代，但这种做法可能会造成模型精度下降，需谨慎考虑。模型可解释性仍是业界难题，对于不可解释模型建议充分测试并监控模型推理结果。

1.8 AI中台模型训练和使用流程

需打造AI中台的数据管理、模型训练、编译优化、模型管理和模型推理等全方位能力。AI中台从数据湖中获取数据，进行数据预处理和标注，将数据送至模型训练模块；由训练模块进行AI模型特征工程、选择合适的算法进行模型的训练；训练完后进行模型的评估，如果模型达不到期望的准确率或消耗的资源过多，还需要进行模型优化（包括超参重新设置、模型压缩），然后再重新进行模型的训练，这当中可能会涉及到重新理解业务需求，获取其他的数据，重新进行数据标注等不同的情况。在模型评估达到要求后，再将模型发布到模型市场，由应用根据需求下载相应的模型进行部署，最后是使用模型进行推理（见图7）。

模型压缩阶段可以进行剪枝、低比特量化、结构压缩等，以便使得模型能适合边缘和终端等资源受限的场景使用。

模型部署阶段可能涉及到云端部署、边缘部署和设备部署的情况，需要具备协同部署能力。

2 AI能力演进

运营商AI能力的演进将是一个长期持续的过程，需结合运营商云网现状、技术成熟度以及运营商云网演进策略等分阶段逐步推进。

具体到未来2~3年内，建议构建并逐渐叠加AI能力来满足运营商自身一体化管控需求和行业需求。

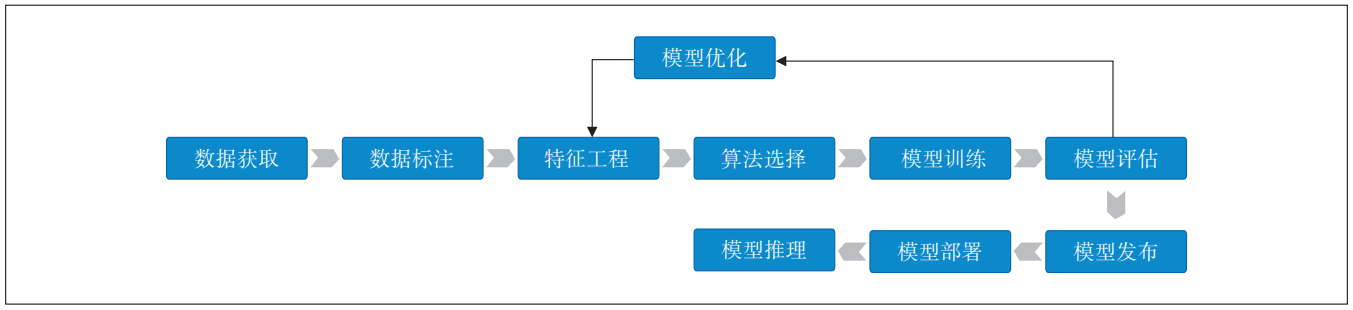

图7 AI模型训练和使用流程

建议分阶段发展如下AI中台能力(具体可根据特定运营商现有AI能力情况做适当调整)。

a)第1阶段:在基础能力方面,构建AI能力,具备机器学习、深度学习训练引擎,推理引擎,端到端支持数据管理、训练、编译优化和推理等基础AI功能;在模型方面,建议构建部分通用AI模型和电信领域AI能力;在数据方面,建议与运营商内部数据打通;在能力共享方面,建议构建模型市场,内部用户可以申请及访问AI能力;在模型运行方面,建议可以做到基于容器/虚机、CPU/GPU运行模型。

b)第2阶段:在基础能力方面,建议构建强化学习和知识图谱能力;在数据方面,建议与网络数据打通;在模型方面,建议进一步丰富通用AI模型和电信领域AI模型,并纳入部分成熟行业AI模型,开始应用流数据更新迭代模型;在能力共享方面,具备可通过系统申请并访问AI能力。

c)第3阶段:在基础能力方面,建议构建安全可信AI框架,具备云边端AI协同部署、AutoML、联邦学习、图神经网络能力;在数据方面,建议与行业第三方伙伴数据打通;在模型构建方面,建议极大地丰富通用AI模型、电信领域AI模型和行业AI模型,能满足大部分场景使用AI模型的要求;在能力共享方面,建议具备外部客户、合作伙伴可申请及访问AI的能力;在模型运行方面,建议具备基于裸机运行容器、基于专用硬件FPGA/ASIC进行模型推理的能力。

3 结束语

随着运营商中台战略的贯彻和中台智能化能力的不断提升,AI将不断帮助运营商对内降本增效,对外提升业务拓展能力,帮助运营商实现数字化转型。

参考文献:

[1] 3GPP. Procedures for the 5G system(5GS):3GPP TS 23.502[S/OL]. [2021-09-21]. ftp://ftp.3gpp.org/Specs/.
[2] 3GPP. Network data analytics services:3GPP TS 29.520[S/OL]. [2021-09-21]. ftp://ftp.3gpp.org/Specs/.
[3] 裴丹,张圣林,裴昶华. 基于机器学习的智能运维[J]. 中国计算机学会通讯,2017,13(12):68-72.
[4] 尤肖虎,张川,谈晓思,等. 基于AI的5G技术——研究方向与范例[J]. 中国科学:信息科学,2018,48(12):1589-1602.
[5] 张嗣宏,左罗. 基于人工智能的网络智能化 发展探讨[J]. 中兴通讯技术,2019,25(2):57-62.
[6] 王海宁,袁祥枫,杨明川. 基于LSTM与传统神经网络的网络流量预测及应用[J]. 移动通信,2019,43(8):37-44.
[7] 钟华. 企业IT架构转型之道:阿里巴巴中台战略思想与架构实战[M]. 北京:机械工业出版社,2017.
[8] 王志军. 中国联通容器化大数据云平台的探索与实践[J]. 信息技术与标准化,2019(5):66-69.
[9] 司炜. 电信运营商IT整合演进策略和规划方案[J]. 电信科学,2018,34(S1):140-149.
[10] 张誌,张延彬,邢庆文,等. 中国移动私有云演进优化探讨[J]. 电信科学,2019(A01):195-201.
[11] 3GPP. Study of enablers for network automation for 5G:3GPP TR 23.791[S/OL]. [2021-09-21]. ftp://ftp.3gpp.org/Specs/.
[12] 尤肖虎,潘志文,高西奇,等. 5G移动通信发展趋势与若干关键技术[J]. 中国科学(信息科学),2014,44(5):551-563.
[13] 张四海,张建华,陈颖,等. B5G系统中基于无线大数据的新兴技术(英文)[J]. 北京邮电大学学报,2018,41(5):52-61.
[14] 王威丽,何小强,唐伦. 5G网络人工智能化的基本框架和关键技术[J]. 中兴通讯技术,2018,24(2):38-42.
[15] 张琰,盛敏,李建东. 大数据驱动的"人工智能"无线网络[J]. 中兴通讯技术,2018,24(2):1-5.
[16] 中国信息通信研究院安全研究所. 人工智能安全白皮书(2018)[R/OL]. [2021-09-21]. https://wenku.baidu.com/view/a3ff59b2b80d6c85ec3a87c24028915f814d8420.html.

作者简介:
陈俊明,毕业于浙江大学,高级工程师,硕士,主要研究方向为5G、AI、算力网络、行业应用等;张洁,毕业于南京师范大学,副教授,硕士,主要研究方向为电信网络、人工智能、计算机应用等;左罗,毕业于电子科技大学,高级工程师,硕士,主要研究方向为5G、AI、云网融合、算力网络等。