

英特尔大数据分析 + AI 平台 助金风慧能打造新能源智能功率预测方案



“金风慧能一直致力于将 IT 前沿技术与新能源的发展相融合, 为用户提供更为高效、敏捷的新能源解决方案。Analytics Zoo 所提供的统一端到端架构, 及其在时序数据分析方面的优势, 不仅使我们结合气象预报数据的多模型组合功率预测方案具有更敏捷的部署效率, 在预测准确率、稳定性方面也能获得较大提升。”

张利
首席架构师
金风慧能

随着清洁能源理念被广泛接受, 以风电、光伏为代表的可再生新能源开始在能源格局中扮演越来越重要的角色。来自国家能源局的数据显示, 仅 2019 年上半年, 全国风电发电量同比增长 11.5%、光伏发电装机量同比增长 20%¹。

但与传统能源不同, 风电、光伏等新能源受环境因素影响很大, 风能、日照的随机性和波动性都会对电网的稳定性和安全性带来严峻挑战。不过随着人工智能 (Artificial Intelligence, AI)、物联网、云计算等新兴技术更多地与新能源产业相融合, 以智能化手段来应对以上挑战, 正成为新能源企业面向未来, 领跑绿色能源发展之路的首选。

矢志于引领新能源数字化、智能化, 致力于新能源物联网系统和解决方案研发的北京金风慧能技术有限公司 (以下简称金风慧能) 就是这条道路上的重要领跑者之一。它正利用深度学习与机器学习的方法, 结合风机级气象预报、风轨迹模拟等气象预报数据, 以多模型组合的方式来构建用于功率预测的全新智能方案。为此, 金风慧能与英特尔开展了紧密的合作, 基于英特尔统一的大数据分析和 AI 平台——Analytics Zoo, 打造了从特征工程搭建、预测影响因子捕捉、多模型组合到定制化策略更新的分布式架构, 并针对预测数据的时序特性进行了有针对性的优化。随后的一系列场站测试中, 双方合力打造的全新方案也得到了验证并收获了良好的应用效果。

金风慧能全新智能功率预测方案实现的优势:

- 与传统功率预测方案相比, 结合气象预报数据的多模型组合功率预测新方案, 能获得更有效的时序数据支撑, 令预测准确率和稳定性更高;
- 英特尔统一的大数据分析和 AI 平台 Analytics Zoo 为新方案提供了统一的端到端架构, 可帮助用户提高系统的开发部署效率和可扩展性, 降低硬件与运维的成本;
- Analytics Zoo 在时序数据分析方面的独特功能和优势, 为新方案提供了更优的预测性能, 可使预测模型与发电设备的实际输出功率实现更好的拟合。

以清洁为重要特点的风电、光伏等可再生能源如今越来越为大众所青睐，但在其生产过程中，风速、风向、日照、气温、气压等环境因素，都会给电力系统带来巨大影响，使发电设备、并入电网都面临运行效率、设备安全等方面的问题。因此，对这些新能源的输出功率开展预测，不仅有利于调度系统合理调整和优化发电计划，改善电网调峰能力，更能减少弃风和弃光率，并降低其运行成本，帮助电力企业提升竞争力。

传统上，电力企业会根据历史资料以及人工经验来实施功率预测，但这一方法往往存在准确率低、波动性大等弊端。随着 IT 技术的不断发展，更多信息化、数字化预测方法正逐渐被运用在功率预测中。尤其是更多 AI 算法的涌现和更强算力的支撑，使得结合海量历史数据样本开展基于 AI 的智能预测方法具备了落地的条件，例如支持向量机 (Support Vector Machine, SVM)、多层感知器 (Multi-Layer Perceptron, MLP) 等，已被实践验证可有效助力电力系统提升运作效率。

但在智能功率预测领域有着深厚积累的金风慧能看来，现有的、基于 AI 的预测方法仍有很大的优化提升空间，其缘由一方面在于单一的深度学习或机器学习方法在准确率和稳定性方面还有待提升；另一方面则在于，预测系统所涉及的一系列模型构建过程，例如特征工程搭建、短期、超短期在线/离线训练等，也对电力企业的 IT 能力提出了挑战。因此，金风慧能亟待引入新的技术，为行业用户提供一个部署更为便捷、预测更为精确的新一代智能功率预测方案。

结合气象预报的多模型组合功率预测方案

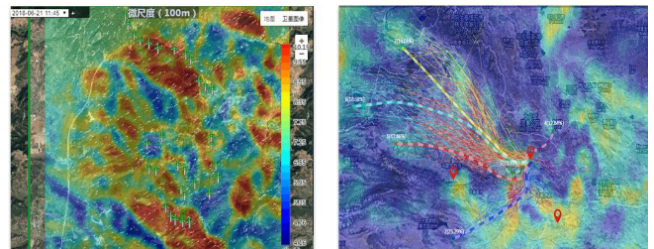
通常，智能功率预测流程可分为边缘数据采集、存储、治理，AI 训练推理等模块。如图一所示，在最左侧，部署在风机、光伏逆变器处的边缘传感器，会把现有的风速、温度、功率、发电量等环境和生产运行数据通过网络传送到数据中心进行存储，而后在数据治理模块中进行有效数据的筛选，并使用相关的深度学习或机器学习框架，将筛选出的数据与历史数据一起进行训练和推理。最终得到的模型，将用于建立在未来一定时间内环境参数与功率、功率与功率、发电量与功率等元素间的映射关系。



图一 传统 AI 功率预测流程

可以看到，传统的 AI 功率预测，主要是通过单一人工智能算法模型，对环境参数、功率、发电量等数据样本来进行训练和推理。这一方法虽然可对较近的时间点 (15-30 分钟内) 有准确的预测，但随着预测时段变长，预测准确率也会随之降低。而对于电力生产所需的功率预测系统而言，最常见的超短期预测也需要系统能预测 4 小时内的功率输出，这意味着预测系统需要在未来 16 个时间点 (每 15 分钟计为一个时间点) 上，都保持出色的预测准确率和稳定性。

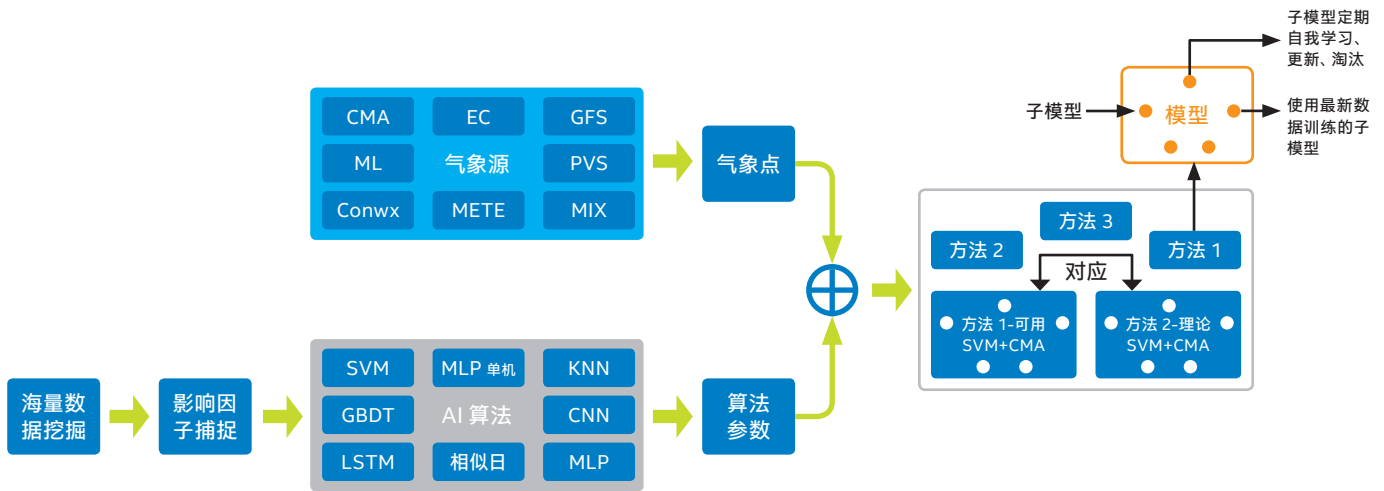
通过研究和实践，金风慧能发现，在既有智能预测系统中，只需导入两项技术即可有效提升预测的准确率和稳定性。首先，是将功率预测与气象预报相结合。由于传统预测方法是一种“从已有功率数据去预测未来功率”的方案，因此其在时间维度上缺乏必要的的数据支撑，而气象预报数据则能有效地弥补这一短板，其时序性数据能令未来各个时间点的预测准确率都保持一致性；另一项技术是采用多模型组合方案，即根据实际需求，选取不同的深度学习或机器学习模型，分别与气象预报数据进行组合，扬长补短，提升系统的预测准确率。



图二 风机级气象预报和风电轨迹模拟

由此，功率预测也对气象预报数据提出了更严格的参数和指标要求，例如更细微的空间分辨率、更为定性及定量的特征分析以及更稳定的预报精度。为此，金风慧能在新方案中与中国气象局 (China Meteorological Administration, CMA)、欧洲气象中心 (European Centre, EC) 等权威机构合作，以集合预报的方式，来保证预报精度的稳定性。如图二所示，新方案中，左图风机级气象预报尺度从常规的 9 公里细化到了 100 米的微尺度，而右图中，对风轨迹的模拟则做到了明确的行进路线定量分析。

结合更高质量的气象预报数据，金风慧能搭建了多模型组合的预测方案。首先，方案通过对海量数据进行挖掘，搭建特征工程，并对与预测相关的影响因子，例如风速、风向、温度等



图三 结合气象预报数据的多模型组合预测方案架构

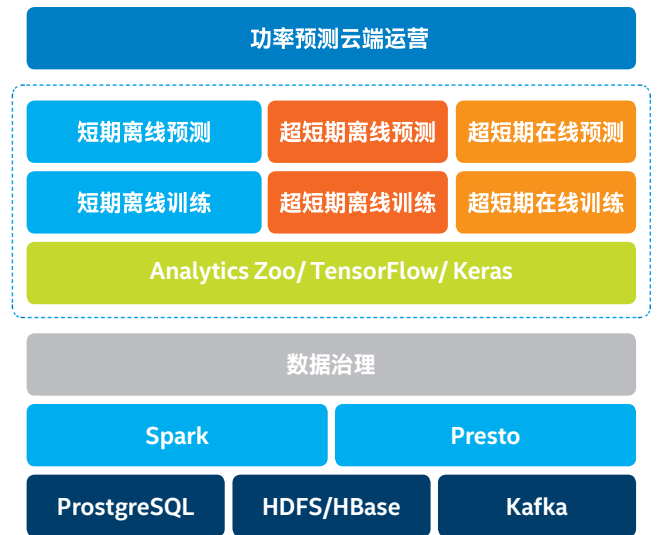
进行捕捉。然后如图三所示，方案将来自不同气象预报源，例如中国气象局、欧洲气象中心的气象预报数据与不同深度学习或机器学习算法的参数相组合，形成诸如 SVM+CMA 的方法。

同时，新方案也可根据需要，形成更多的组合方法。各种组合方法中的算法模型，也能够不断使用最新数据对其子模型进行训练，从而完成子模型的定期自我学习、更新和淘汰，以不断迭代的方式提升预测系统的准确率。

Analytics Zoo 助力分布式功率预测架构搭建

金风慧能新方案的主要优势是与气象预报数据进行了结合。众所周知，气象预报数据是一种典型的时序数据，其数值会随着时间的变化而发生变化。因此，选择更适用于时序数据处理的系统架构，无疑可进一步提升预测系统的准确性和稳定性，而英特尔统一的大数据分析和 AI 平台——Analytics Zoo，不仅可为方案提供统一的端到端分布式方案，帮助用户提升系统的开发部署效率和可扩展性，同时其在时序数据分析方面的独特功能和优势也能为新方案提供更强助力。

作为英特尔在大数据分析与 AI 领域的融合型创新成果，Analytics Zoo 如图四所示，可帮助金风慧能将新方案中的 Spark、TensorFlow、Keras 及其它软件和框架无缝集成到同一管道中。这一方法有助于金风慧能将数据存储、数据处理以及训练推理的流水线整合到统一的基础设施上，来大幅提升新方案的部署效率、资源利用率和可扩展性，并减少用于硬件管理及系统运维的成本。



图四 基于 Analytics Zoo 的分布式功率预测架构

同时，Analytics Zoo 还能卓有成效地将英特尔提供的众多底层软件加速库，如英特尔® 数学核心函数库 (Intel® Math Kernel Library, 英特尔® MKL)、面向深度神经网络的英特尔® 数学核心函数库 (Intel® Math Kernel Library for Deep Neural Network, 英特尔® MKL-DNN) 等，应用到上层功率预测方案的优化中去；并可将 TensorFlow、Keras 模型透明地扩展到大数据集群，使用户能更方便地在训练或推理方案中采用分布式架构，以进一步提升性能表现。

在提供统一的端到端平台架构之外，Analytics Zoo 给予新方案的优势，还在于其提供了一系列与时序数据分析相关的功能和特性。首先，Analytics Zoo 对于不同时序分析应用，如时序

预测、异常检测、时序表征学习、时序聚类，都可提供完整的解决方案，便于金风慧能在新方案中构建更多的预测方法组合；其次，Analytics Zoo 针对时序数据预置了丰富的功能组件，包括：

- 功率预测常见的深度学习和机器学习模型：LSTM、Encoder-Decoder、MTNet、ARIMA 等；
- 功率预测中常用的数据预处理和特征工程：Datetime features、Time diff、Log-transform、Rolling window 等；
- 功率预测中普遍的异常探测方法：Percentile、Distribution-based、Uncertainty based、Autoencoder 等。

除了以上组件，Analytics Zoo 还可为新方案提供 AutoML (自动机器学习) 方法，使之能够进行自动化特征选择、模型选择和超参调优等，令预测模型得以更好地拟合发电设备输出功率的变化周期。

为验证基于 Analytics Zoo 的全新分布式功率预测方案在实际运行中的表现，金风慧能与英特尔一起，在全国多个光伏测试场站进行了实地测试。验证方案以月为周期，在每一个测试的光伏场中，在单小时内使用 30,000 条记录对 LSTMNet 模型进行 5,000 次迭代优化，并在 50 毫秒内获得未来 2 小时的功率预测数据。验证结果如图五所示，新方案在预测准确率上超越了原有方案的 59%，达到了 79.41%²。



¹ 数据援引自国家能源局数据统计报告：http://www.nea.gov.cn/2019-07/26/c_138259422.htm, http://www.nea.gov.cn/2019-08/23/c_138330885.htm

² 测试配置为：处理器：英特尔® 至强® 金牌 6130 处理器；内存：192GB DDR4 2666MHz；操作系统：CentOS 7.6；Spark 版本：2.4.3，更多细节请与金风慧能联系。

英特尔并不控制或审计第三方数据。请您审查该内容，咨询其他来源，并确认提及数据是否准确。

描述的成本降低情景均旨在特定情况和配置中举例说明特定英特尔产品如何影响未来成本并提供成本节约。情况均不同。英特尔不保证任何成本或成本降低。

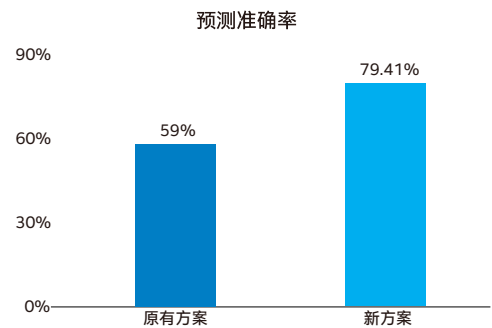
英特尔技术特性和优势取决于系统配置，并可能需要支持的硬件、软件或服务得以激活。产品性能会基于系统配置有所变化。没有任何产品或组件是绝对安全的。更多信息请从原始设备制造商或零售商处获得，或请见 intel.com。

没有任何产品或组件是绝对安全的。

英特尔技术可能需要支持的硬件、软件或服务得以激活。请从原始设备制造商或零售商处获得更多信息。

英特尔、英特尔标识以及其他英特尔商标是英特尔公司或其子公司在美国和/或其他国家的商标。

©英特尔公司版权所有



图五 金风慧能功率预测新旧方案对比

展望：从引领创新到加速实践

金风慧能与英特尔通过协作，不仅开发出了更为完善、高效和精确的全新智能功率预测方案，还已着手将它用于实践。来自金风慧能的数据统计表明，在风电场中引入这种更为高效的 AI 预测方案，能够帮助电力企业显著提升发电效率，为践行绿色环保的新能源理念提供强有力的支撑。

未来，双方还计划在新能源智能功率预测方案上开展更多、更深入的合作。一方面，金风慧能希望进一步基于 Analytics Zoo 搭建完整的数据治理平台，优化数据质量；另一方面，双方还计划将新方案的验证工作扩大到 100 个光伏场站，并拓展到风电功率预测场景，从而以更多的数据、更优的模型与方案，向用户提供高质量、高可用的新能源智能功率预测系统，为 IT 技术与绿色能源的融合与发展持续提供更多、更佳的创新与实践。