

微博深度学习平台架构和实践

原创：黄波，何沧平 人工智能头条 2017-10-12



作者 | 黄波，何沧平

责编 | 何永灿

随着人工神经网络算法的成熟、GPU计算能力的提升，深度学习在众多领域都取得了重大突破。本文介绍了微博引入深度学习和搭建深度学习平台的经验，特别是机器学习 workflow、控制中心、深度学习模型训练集群、模型在线预测服务等核心部分的设计、架构经验。微博深度学习平台极大地提升了深度学习开发效率和业务迭代速度，提高了深度学习模型效果和业务效果。

深度学习平台介绍

人工智能和深度学习

人工智能为机器赋予人的智能。随着计算机计算能力越来越强，在重复性劳动和数学计算方面很快超过了人类。然而，一些人类通过直觉可以很快解决的问题，例如自然语言理解、图像识别、语音识别等，长期以来很难通过计算机解决。随着人工神经网络算法的成熟、GPU计算能力的提升，深度学习在这些领域也取得了重大的突破，甚至已经超越人类。深度学习大大拓展了人工智能的领域范围。

深度学习框架

深度学习框架是进行深度学习的工具。简单来说，一套深度学习框架就是一套积木，各个组件就是某个模型或算法；开发者通过简单设计和组装就能获得自己的一套方案。深度学习框架的出现降低了深度学习门槛。开发者不需要编写复杂的神经网络代码，只需要根据自己的数据集，使用已有模型通过简单配置训练出参数。

TensorFlow、Caffe和MXNet是三大主流的深度学习开源框架：TensorFlow的优势是社区最活跃，开源算法和模型最丰富；Caffe则是经典的图形领域框架，使用简单，在科研领域占有重要地位；MXNet在分布式性能上表现优异。PaddlePaddle、鲲鹏、Angel则是百度、阿里、腾讯分别推出的分布式计算框架。

2015年底，Google开源了TensorFlow深度学习框架，可以让开发者方便地组合CNN、RNN等模块实现复杂的神经网络模型。TensorFlow是一个采用数据流图（data flow graphs），用于数值计算的开源软件库。

2016年，百度开源了PaddlePaddle（PArallel Distributed Deep LEarning 并行分布式深度学习）深度学习框架。PaddlePaddle具有易用，高效，灵活和可伸缩等特点，为百度内部多项产品提供深度学习算法支持。

深度学习平台

深度学习框架主要提供神经网络模型实现，用于进行模型训练。模型训练只是机器学习和深度学习中的一环，除此之外还有数据输入、数据处理、模型预测、业务应用等重要环节。深度学习平台就是整合深度学习各环节，为开发者提供一体化服务的平台。深度学习平台能够加快深度学习的开发速度，缩减迭代周期；同时，深度学习平台能够将计算能力、模型开发能力共享，提升开发效率和业务效果，也能够将资源合理调度，提高资源利用率。

腾讯深度学习平台DI-X

腾讯深度学习平台DI-X于2017年3月发布。DI-X基于腾讯云的大数据存储与处理能力来提供一站式的机器学习和深度学习服务。DI-X支持TensorFlow、Caffe以及Torch等三大深度学习框架，主要基于腾讯云的GPU计算平台。DI-X的设计理念是打造一个一站式的机器学习平台，集开发、调试、训练、预测、部署于一体，让算法科学家和数据科学家，无须关注机器学习（尤其是深度学习）的底层工程繁琐的细节和资源，专注于模型和算法调优。

DI-X在腾讯内部使用了一年，其主要用于游戏流失率预测、用户标签传播以及广告点击行为预测等。

阿里机器学习平台PAI

阿里机器学习平台PAI1.0于2015年发布，包括数据处理以及基础的回归、分类、聚类算法。阿里机器学习平台PAI2.0于2017年3月发布，配备了更丰富的算法库、更大规模的数据训练和全面兼容开

源的平台化产品。深度学习是阿里机器学习平台PAI2.0的重要功能，支持TensorFlow、Caffe、MXNet框架，这些框架与开源接口兼容。在数据源方面，PAI2.0支持非结构化、结构化等各种数据源；在计算资源方面，支持CPU、GPU、FPGA等异构计算资源；在工作流方面，支持模型训练和预测一体化。

PAI已经在阿里巴巴内部使用了2年。基于该平台，在淘宝搜索中，搜索结果会基于商品和用户的特征进行排序。

百度深度学习平台

百度深度学习平台是一个面向海量数据的深度学习平台，基于PaddlePaddle和TensorFlow开源计算框架，支持GPU运算，为深度学习技术的研发和应用提供可靠性高、扩展灵活的云端托管服务。通过百度深度学习平台，不仅可以轻松训练神经网络，实现情感分析、机器翻译、图像识别，也可以利用百度云存储和虚拟化产品直接将模型部署至应用环境。

微博深度学习平台设计

微博在Feed CTR、反垃圾、图片分类、明星识别、视频推荐、广告等业务上广泛使用深度学习技术，同时广泛使用TensorFlow、Caffe、Keras、MXNet等深度学习框架。为了融合各个深度学习框架，有效利用CPU和GPU资源，充分利用大数据、分布式存储、分布式计算服务，微博设计开发了微博深度学习平台。

微博深度学习平台支持如下特性：

- **方便易用：**支持数据输入、数据处理、模型训练、模型预测等工作流，可以通过简单配置就能完成复杂机器学习和深度学习任务。特别是针对深度学习，仅需选择框架类型和计算资源规模，就能模型训练。
- **灵活扩展：**支持通用的机器学习算法和模型，以及用户自定义的算法和模型。
- **多种深度学习框架：**目前支持TensorFlow、Caffe等多种主流深度学习框架，并进行了针对性优化。
- **异构计算：**支持GPU和CPU进行模型训练，提高模型训练的效率。
- **资源管理：**支持用户管理、资源共享、作业调度、故障恢复等功能。
- **模型预测：**支持一键部署深度学习模型在线预测服务。

微博深度学习平台架构和实践

微博深度学习平台是微博机器学习平台的重要组成部分，除继承微博机器学习平台的特性和功能以外，支持TensorFlow、Caffe等多种主流深度学习框架，支持GPU等高性能计算集群。微博深度学习平台架构如图1所示。

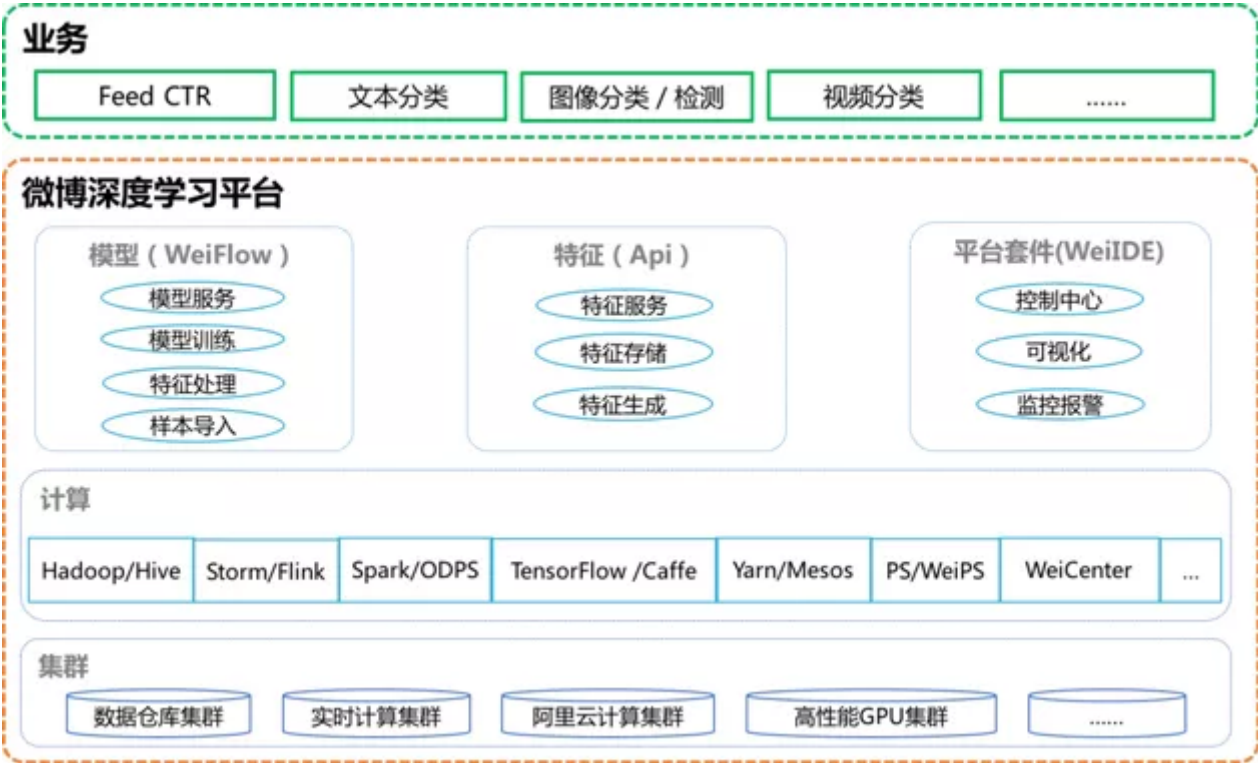


图1 微博深度学习平台架构

下面将以机器学习 workflow、控制中心、深度学习模型训练集群、模型在线预测服务等典型模块为例，介绍微博深度学习平台的实践。

机器学习 workflow WeiFlow

微博深度学习和机器学习 workflow 中，原始数据收集、数据处理、特征工程、样本生成、模型评估等流程占据了大量的时间和精力。为了能够高效地端到端进行深度学习和机器学习的开发，我们引入了微博机器学习 workflow 框架 WeiFlow。

WeiFlow 的设计初衷就是将微博机器学习流的开发简单化、傻瓜化，让业务开发人员从纷繁复杂的数据处理、特征工程、模型工程中解脱出来，将宝贵的时间和精力投入到业务场景的开发和优化当中，彻底解放业务人员的生产力，大幅提升开发效率。

WeiFlow的诞生源自于微博机器学习的业务需求。在微博的机器学习 workflows 中（如图2所示），多种数据流经过实时数据处理，存储至特征工程并生成离线的原始样本。在离线系统，对原始样本进行各式各样的数据处理、特征处理、特征映射，从而生成训练样本；业务人员根据实际业务场景（排序、推荐），选择不同的算法模型，进行模型训练、预测、测试和评估；待模型迭代满足要求后，通过自动部署将模型文件和映射规则部署到线上。线上系统根据模型文件和映射规则，从特征工程中拉取相关特征，根据映射规则进行预处理，生成可用于预测的样本格式，进行线上实时预测，最终将预测结果（用户对微博内容的兴趣程度）输出，供线上服务调用。



图2 微博机器学习工作流

为了应对微博多样的计算环境，WeiFlow采用了双层的DAG任务流设计，如图3所示。外层的DAG由不同的Node构成，每一个Node是一个内层的DAG，具备独立的执行环境，即上文提及的Spark、TensorFlow、Hive、Storm、Flink等计算引擎。

图3 WeiFlow双层DAG任务流设计

外层DAG设计的初衷是利用各个计算引擎的特长，同时解决各个计算引擎间的依赖关系和数据传输问题。内层的DAG，利用引擎的特性与优化机制，实现不同的抽象作为DAG中计算模块之间数据交互的载体。

在使用方面，业务人员根据事先约定好的规范和格式，将双层DAG的计算逻辑定义在XML配置文件中。依据用户在XML指定的依赖关系和处理模块，WeiFlow自动生成DAG任务流图，并在运行时阶段调用处理模块的实现来完成用户指定的任务流。通过在XML配置文件中将所需计算模块按照依赖关系堆叠，即可以搭积木的方式完成配置化、模块化的流水线作业开发。

控制中心WeiCenter

控制中心WeiCenter的目标就是简单、方便、易用，让大家便利地使用微博深度学习平台。下面将介绍控制中心的作业管理、数据管理和调度管理等部分。

- **作业管理：**我们在进行深度学习、大规模机器学习、实时处理的过程中，由于需要各种不同框架的配合使用共同完成一个任务，比如TensorFlow适合进行高性能学习、Spark适合大规模亿维特征训练、Storm或者Flink适合实时特征生成以及实时模型生成等，将这些结合到一起才能完成从离线训练到线上实时预测。以前这需要开发者去学习各种框架复杂的底层开发，现在通过控制中心选择不同的作业类型，可以方便地生成各种类型的作业任务。用户只需要在可视化UI上进行作业类型选择、数据源选择、输出目的地选择或者使用WeiFlow进行编程，就能生成一个高大上的深度学习或机器学习作业。
- **数据管理：**当大数据的数据量，每天按P级增长，使用人员每天上百人时，数据管理就显得尤为重要。如果模型训练的集群和数据所在的集群，不是同一个集群，如何高效地将数据同步到模型训练的集群是一个难点。并且在完成模型训练后，能自动根据训练结果作出评估，对训练数据进行删除。由于使用集群的开发人员素质不齐，你会发现总是有很多冗余数据没删除，而且总有无用数据生成，这个时候需要一个统一的数据管理平台，去约束大家生成数据的同时删除数据，去各个平台上探测长时间无访问的数据并进行确认清理。
- **调度管理：**作业有多种分类，按重要程度分：高、中、低；按占用资源量分：占用多、占用一般、占用少；按调度器分：Yarn、Mesos、Kubernetes等。Spark、Hadoop利用Yarn调度解决了优先级高的作业和资源占用多作业之间的矛盾；TensorFlow利用成熟的Kubernetes或Mesos调度TensorFlow节点进行GPU集群化任务管理；普通离线作业和服务部署利用Mesos进行资源调度。控制中心集成了多种调度器，利用各种成熟的解决方案，简化了作业负责调度这一难题。

总之，控制中心负责用户权限控制、作业图依赖管理、数据依赖管理等，调度服务负责具体的作业执行、资源抽象、资源管理。控制中心和调度服务如图4所示。




图4 控制中心和调度服务

深度学习模型训练集群

微博深度学习训练集群与传统HPC集群有重大区别，分别体现在计算服务器选型、分布式训练、网络设备、存储系统、作业调度系统。

- **单机多GPU卡**：深度学习模型训练大部分情况下单机运算，且几乎完全依靠GPU，因此选用能挂载2/4/8块GPU的服务器，尽量提高单机运算能力。
- **分布式训练**：如果训练时间长或者样本规模大，超过单台服务器能力时，需要支持分布式训练。以TensorFlow分布式运行方式为例进行说明，如图5所示。一个TensorFlow分布式程序对应一个抽象的集群，集群（cluster）由工作节点（worker）和参数服务器（parameter server）组成。工作节点（worker）承担矩阵乘、向量加等具体计算任务，计算出相应参数（weight和bias），并把参数汇总到参数服务器；参数服务器（parameter server）把从众多工作节点收集参数汇总并计算，并传递给相应工作节点，由工作节点进行下一轮计算，如此循环往复。

图5 TensorFlow分布式架构

- **万兆以太网网络**：参数更新过程中，通信粒度大，而且允许异步通信，对延时没有严格要求。因此，训练集群没有选用HPC集群必备的InfiniBand或Omini-Path低延时网络设备，而是选用普通的以太网设备。
- **HDFS分布式文件系统**：TensorFlow分布式工作节点读取训练样本文件时，不同工作节点读取的数据段不交叉，训练过程中也不会交换样本数据。写出模型文件也只有某一个工作节点负责，不涉及工作节点间的数据交换。因此，深度学习训练作业不要求HPC机群中常见的并行文件系统，只要是一个能被所有工作节点同时访问文件系统就可以。实际上，微博深度学习平台采用HDFS，不但满足要求，而且方便与其它业务共享数据。
- **定制的作业调度系统**：TensorFlow分布式参数服务器进程不会自动结束，需要手动杀死，而HPC应用中的MPI进程同时开始同时结束。设计作业调度方案时必须考虑这个特点，使之能够在所有工作节点都运行结束后自动杀死参数服务器进程。

模型在线预测服务WeiServing

模型在线预测服务是深度学习平台的一个重要功能。由于微博业务场景需求，模型在线预测服务并发量大，对延时、可用性要求极高。考虑到这些业务需求以及服务本身以后的高扩展性，微博分布式模型在线预测服务WeiServing的架构如图6所示。

图6 WeiServing-微博分布式模型在线预测服务架构

- **特征处理多样化**：模型在线预测服务首先要解决的问题是，将在线的原始特征数据，映射成模型可以处理的数据格式。基于大量的业务模型实践与调优，微博机器学习 workflow 框架 WeiFlow 抽象出了一套特征处理函数，来提升开发效率和业务效果。WeiServing 与 WeiFlow 在特征处理方面一脉相承，支持一系列特征处理函数，包括 `piecewise`、`pickcat`、`descartes`、`combinehash` 等映射函数，对特征进行归一化、离散化、ID 化、组合等特征处理。
- **多模型多版本支持**：由于微博业务场景多种多样，不同的业务场景对模型与特征有不同的需求，WeiServing 支持同一个集群为多个业务提供服务，通过 `docker+k8s` 进行资源隔离与负载均衡。在相同特征情况下，可以选择不同的模型算法进行处理。另外，对于同一个模型，WeiServing 支持在线升级与多版本同时在线，为业务灰度测试提供可能。所有的差异化都被映射到配置文件中，通过简单的配置来完成线上模型的转换。
- **分布式服务支持**：为了应对大规模模型服务与在线机器学习，WeiServing 参考通用的参数服务器解决方案，实现了 WeiParam 分布式服务架构，除了支持传统的 PS 功能之外，WeiParam 针对在线服务需求，通过分布式调度系统，提供多副本、高可用、高性能的系统机制。
- **多源支持**：对于普通离线学习，模型会导出到文件中，WeiServing 通过 ModelManager 模块管理模型加载，支持本地存储与分布式存储。同时，WeiServing 为支持在线机器学习，提供对实时流接口对接，在线训练的模型参数可以实时推送到 WeiParam 中，为线上提供服务。

总结

本文介绍了深度学习框架和平台的概念和特征，基于微博深度学习平台深入探讨了深度学习平台的设计思考和技术架构。机器学习 workflow 和控制中心是我们在规范机器学习 workflow 的设计成果，系统化的标准流程能极大提升机器学习开发效率和业务迭代速度。

深度学习模型训练集群和模型在线预测服务是我们在深度学习模型训练、模型预测的集群化、服务化方面的系统产出，是保障模型效果和业务效果的基础。希望上述介绍能给大家带来思考和帮助。

作者简介：

黄波，微博机器学习平台技术负责人、资深架构师，负责微博机器学习平台和 Feed 算法系统。毕业于中科院计算所，熟悉 Feed 系统、推荐系统、分布式计算系统。

何沧平，微博深度学习平台架构师，主要负责深度学习算法在微博 Feed 个性化排序、个性化推荐等业务中的应用。毕业于中科院数学院，熟悉高性能计算，著有图书《OpenACC 并行编程实战》。

本文为《程序员》原创文章，未经允许不得转载，更多精彩文章点击「[阅读原文](#)」订阅《程序员》

SDCC 2017“人工智能技术实战线上峰会”将在 CSDN 学院以直播互动的方式举行。

作为 SDCC 系列技术峰会的一部分，来自 [阿里巴巴](#)、[微软](#)、[商汤科技](#)、[第四范式](#)、[微博](#)、[出门问问](#)、[菱歌科技的 AI 专家](#)，将针对机器学习平台、系统架构、对话机器人、芯片、推荐系统、Keras、分布式系统、NLP 等热点话题进行分享。先行者们正在关注哪些关键技术？如何从理论跨越到企业创新实践？你将从本次峰会找到答案。[每个演讲时段均设有答疑交流环节，与会者和讲师可零距离互动。](#)

阅读原文