

UNIS DataEngine ADE 应用开发引擎

背景介绍

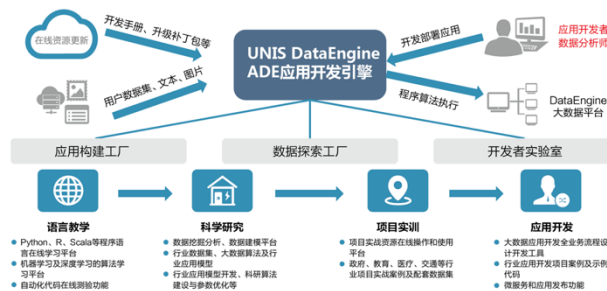
随着大数据战略的快速发展，以及行业分析报告来看，大数据底层技术已进入实际落地阶段，基于 Hadoop 生态系统、MPP 分布式数据库的大数据底层平台已基本能够满足企业对海量数据的采集存储查询的需求。随着数据的积累，数据拥有者越来越渴望对数据价值的追求，人们希望通过应用开发手段能够在海量数据集中挖掘其中潜在数据规律、关联规则，从传统基于存储统计的辅助决策向智能决策进行转变。但大数据应用开发，和传统 web 应用开发又存在较大差异。

从行业调研来看，大数据应用开发在实际上面临常见的困难：



产品定位于使命

UNIS DataEngine ADE (Application Develop Engine) 应用开发引擎是聚焦大数据应用开发及数据分析挖掘的平台及服务。整个系统基于底层分布式批处理框架、流数据实时框架、分布式数据库以及大规模机器学习算法库等大数据相关技术，提供各类应用开发套件。提高企业智能应用开发的能力和效率、简化复杂机器学习算法的使用成本，从而帮助企业实现数据驱动的商业模式。



UNIS DataEngine ADE 应用开发引擎概述

ADE 产品系统架构

UNIS DataEngine ADE 应用开发引擎系统结构如下图所示：

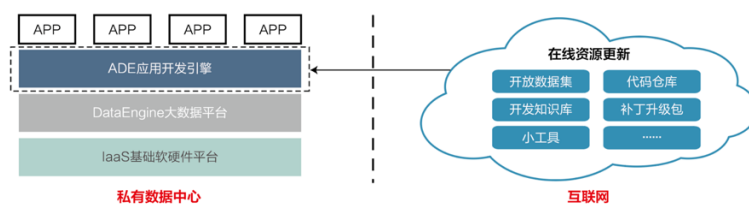


图 DataEngine ADE 应用开发引擎架构定位

ADE 产品构架于 DataEngine 大数据平台之上，对上向应用开发者提供一站式开发能力。并通过在线资源更新，可实时获取开发数据集、开发知识库、小工具、补丁包等资源。

➤ ADE 产品详细功能

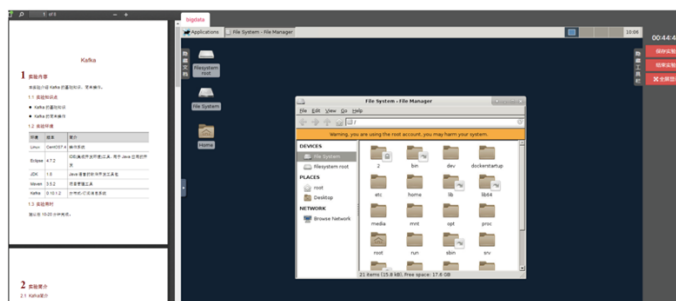
◆ 开发者实验室

开发者实验室主要是面向 IT 初学者提供在线学习技术的平台，实现用户零门槛的学习，不同课程提供不同的虚拟化环境，用户可以边学习边操作，提供用户一站式的学习。对学生用户，开发者实验室提供上机实验功能，对于教师，开发者实验室提供上机实验和实验管理功能。实验课程内容有大数据技术原理与应用、Spark 编程基础、Python 数据分析、Python 可视化、SQL 基础、JAVA 基础等内容。



上机列表主页面

点击对应实验列表中的[]按钮，开始学习该实验，实验界面如图，实验界面分为左中右三部分，左边展示该实验的实验文档；中间是上机环境，用户可以边看实验文档边在环境中操作；右侧是工具栏。



◆ 应用构建工厂

应用构建工厂是 ADE 平台的重要功能，采用图形化、向导式进行大数据应用构建。

应用构建的主要目的：

- ◎帮助用户快速搭建一个 Web 项目，方便用户进行二次开发。
- ◎减少用户开发工作量。例如，用户通过对关系型数据库组件进行配置，生成的工程中包含了对数据表的增、删、改、查基本操作，可以极大地减少用户的开发工作量。

Code 构建列表详细记录了用户构建项目的历史信息，可以对其进行追溯和维护，极大地方便了用户对历史工程的管理。如图所示：



图 1-1 Code 建

◆ 应用场景列表

用户点击[Code 构建]界面的<新增>按钮，进入[应用场景]界面，如图所示。



发现数据模块

应用场景包含了各式各样的功能模块列表。用户通过不同的功能模块，可以快速的开发不同的工程。

应用场景界面分为统计报表类应用模板、通用脚手架工程模板两种。

⊙用户通过统计报表类应用模板，可以快速地将数据存储到数据库，并生成一个可视化的可展示数据详情的 Web 工程。

⊙用户在通用脚手架工程模板中，可通过拖拽组件，配置组件信息，快速开发可运行的 Web 工程。

在应用构建的左边菜单中，有发现数据模块，其中包括文件数据、数据库数据、网络数据、消息队列、小文件存储、全文检索等具体场景。

◆ 探索数据模块

探索数据模块包括算法教学、回归分类、舆情分析、关联分析、个性推荐、社交网络等场景，其中某些场景中会有更为详尽的分类，如在算法教学中，按照算法的功能，分为分类、聚类、回归、遗传算法等。



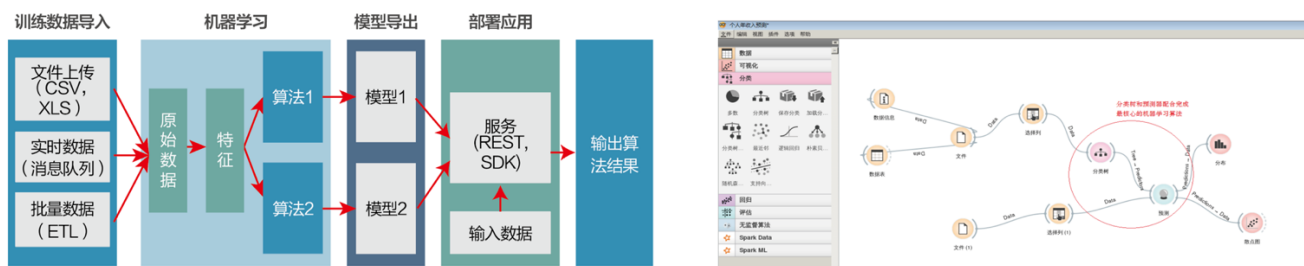
探究数据模块

◆ 数据探索工厂

数据探索工厂提供了标准的数据挖掘模型生成流程，包括数据预处理、特征提取、模型训练、评估评测、导出部署五大关键步骤。

大数据分析的核心是高校精准的数据挖掘算法库，ADE 提供了基于分布式计算框架编写的高性能大规模机器算法库，从而保证对行业全量数据的模型构建，避免了传统分析工具仅能利用小数据样本进行分析挖掘的弊端。

系统在用户体验方面，提供了全流程可视化操作 WebUI，同时支持模型迭代训练过程的实时可视化展现，能够对模型训练过程实现全透明管理监控，辅助数据科学家/数据分析师构建高性能精准的挖掘模型。



◆ 应用部署管理

提供一键式应用部署，并可对应用进行增量部署。可设置后台大数据分析定时任务。

通过应用监控组件，提供应用参数配置、接口调用检测、辅助排障、性能看板等功能。

可通过插件化应用补丁和升级系统，进行应用日常维护。

任务名称	任务开始时间	任务结束时间	任务状态	操作
16	2017-05-22 17:15:07	2017-05-22 17:15:22	2017-05-22 17:16:04	成功
2	2017-05-22 17:16:07	2017-05-22 17:16:22	2017-05-22 17:16:04	成功
7	2017-05-22 17:45:07	2017-05-22 17:50:22	2017-05-22 17:50:42	成功
25	2017-05-22 18:00:07	2017-05-22 18:05:22	2017-05-22 18:05:42	成功
26	2017-05-22 18:15:07	2017-05-22 18:20:22	2017-05-22 18:20:42	成功
17	2017-05-22 18:30:07	2017-05-22 18:35:22	2017-05-22 18:35:42	成功
8	2017-05-22 18:45:07	2017-05-22 18:50:22	2017-05-22 18:50:42	成功
3	2017-05-22 19:00:07	2017-05-22 19:05:22	2017-05-22 19:05:42	成功
12	2017-05-22 19:15:07	2017-05-22 19:20:22	2017-05-22 19:20:42	成功
27	2017-05-22 19:30:07	2017-05-22 19:35:22	2017-05-22 19:35:42	成功

◆ 开放数据集

海量数据已经成为各大互联网企业面临的最大的问题，如何处理海量数据，提供更好的解决方案，是目前热门的话题。那么怎样获取海量数据将是面临的另外一个难题。开放数据集部分，我们按照行业划分，将数据集划分为不同情况，目前划分为统计、互联网、科研、电商等。可以通过点击<样例数据>按钮就行样例数据下载，查看该数据集的样例部分，如果能够满足用户需要，则可以继续点击<下载>按钮，下载完整数据集以作为处理和分析。



开放数据集

◆ 知识库

知识库，你身边的技术百科全书。在 ADE 的知识库模块中，可以下载大量的技术文档来满足不同的需求。如开发指南、用户手册、API 手册、实践总结等等。点击<下载>按钮即可完成知识库文档的下载，如图所示：



➤ ADE 产品核心优势

◆ 大数据应用 快速构建

- ⊙ 自动化生成软件开发工程和框架代码，减少开发者编码量。
- ⊙ 示例代码、用户故事、开发手册等丰富知识库。

◆ 精准分析 深度挖掘

- ⊙ 全流程可视化模型构建，降低数据挖掘、机器学习落地门槛。
- ⊙ 丰富的大规模机器学习算法库，挖掘分析准确度高。
- ⊙ 开发数据集、软件仓库、各类开发过程涉及小工具、补丁升级包。
- ⊙ 创新灵活的生态体系。

➤ ADE 产品适合的行业场景

软件开发商	高校	政府	其他
快速具备大数据开发能力 创新思维、用户故事 从传统分析到智能挖掘 在线丰富资源库	实训课程，编程开发 科研环境，建模工具 创新大赛，人才培养	零基础数据应用开发 数据挖掘驱动智能服务、智能监管、 犯罪预测、规划设计等	海量数据 + 软件开发技术 + 算法技术 = 智能应用

北京紫光恒越网络科技有限公司

<http://www.unishy.com>

UNIS

北京基地
北京市海淀区中关村东路 1 号院 2 号楼 402 室
邮编：100084
电话：010-62166890
传真：010-51652020-116
版本：

客户服务热线
400-910-9998

Copyright ©2012 北京紫光恒越网络科技有限公司 保留一切权利
免责声明：虽然 UNIS 试图在本资料中提供准确的信息，但不保证资料的内容不含有技术性误差或印刷性错误，为此 UNIS 对本资料中的不准确不承担任何责任。
UNIS 保留在没有通知或提示的情况下对本资料的内容进行修改的权利。