

-

KuAI-产品白皮书

A

gile

BI

平台

-

产品白皮书

目录

[1. 产品概述 2](#_Toc67073374)

[2. 产品价值 2](#_Toc67073375)

[2.1. 一站式：覆盖机器学习建模全生命周期流程，高效开发高效管理 2](#_Toc67073376)

[2.2. 易用性：多重模型开发方式，满足不同基础用户和不同场景的开发需求 2](#_Toc67073377)

[2.3. 专业性：为金融机构提供更高效的、更可信的模型开发方案 3](#_Toc67073378)

[3. 产品架构 3](#_Toc67073379)

[4. 产品功能 3](#_Toc67073380)

[5. 产品优势 5](#_Toc67073381)

[5.1. 智能化开发，大幅提升机器学习建模效率 5](#_Toc67073382)

[5.2. 可解释模型，符合金融机构合规运营要求 5](#_Toc67073383)

[5.3. 高度适应性，满足不同业务场景的使用需求 5](#_Toc67073384)

[6. 应用场景 5](#_Toc67073385)

[7. 典型案例 5](#_Toc67073386)

# 产品概述

KuAI（智能AI开发平台）是为AI从业者打造的的机器学习全流程平台，旨在面向金融行业大数据探索及建模诉求，通过输出京东科技在AI建模方面的经验产品、技术，一站式地提供集数据探索、特征工程、模型实验、模型训练、模型管理、模型部署、模型监控、模型服务的全链路智能模型构建及管理服务，助力金融机构快速构建业务模型，加速智能化发展，提升客户体验。

我们致力于提升建模效率和降低建模门槛，对于专业建模人员、数据分析人员和业务人员均可以快速上手，对IT人员也可以极大减少IT支持工作，提升调度和数据运作效率。KuAI平台提供编码开发、可视化开发、自动学习三种模型开发训练方式，并封装丰富的风控及营销场景组件和模板，用户可针对业务场景快速实现业务模型，同时支持数据探索、数据预处理、特征工程、模型导入、模型管理、模型发布、模型测试、模型监控等，覆盖机器学习项目从准备数据到模型上线运行全生命周期。

# 产品价值

## 一站式：覆盖机器学习建模全生命周期流程，高效开发高效管理

系统涵盖了数据探索、模型实验、模型训练、模型管理、模型部署、模型监控等功能，支持机器学习项目从准备数据到模型上线运行全生命周期。 一站式的界面操作，能够引导用户快速构建高质量模型并进行模型的部署和使用，同时支持模型的监控，方便用户精、准、快地进行模型的更新迭代，以更好地适应业务要求。

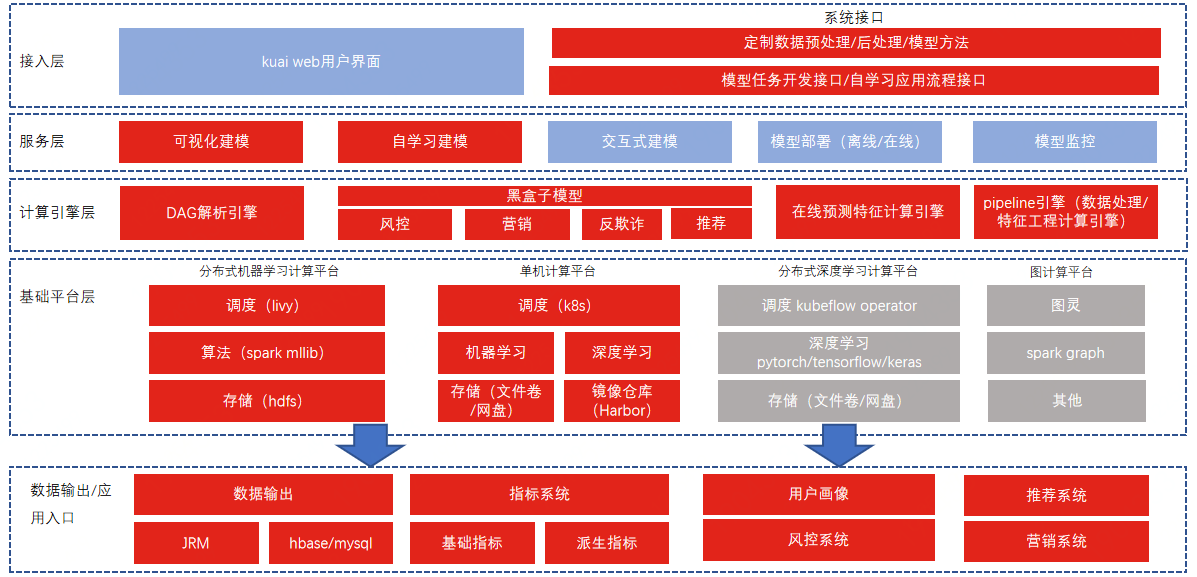
## 易用性：多重模型开发方式，满足不同基础用户和不同场景的开发需求

系统支持自动学习、可视化开发和 编码开发三种方式。自动学习，通过定义任务类型和优化目标，上传需要训练的数据，进行一键建模；可视化开发，通过一站式图形化操作，开发模型、训练模型、预测模型、评估模型、调度模型；编码开发，基于开源 Jupyter Notebook 深度定制，用户可方便的加载数据源、进行模型开发调试、模型数据持久化。简单易用的界面操作不仅能保障普通数据分析师和业务人员能够迅速上手机器学习，又能为资深数据科学家提供高效率的交互式体验并缩短高质量模型开发的周期。

## 专业性：为金融机构提供更高效的、更可信的模型开发方案

针对金融的风控、营销场景，封装丰富的数据探索、数据处理、模型训练组件，以及不同场景的最佳实践pipeline和自动学习模板，用户可以快速构建金融场景模型，无需关注复杂的组成部分。同时考虑金融机构对于模型可溯源和可解释性的高度要求，在数据和模型方面提供可信的报告支撑，助力合规运营。

# 产品架构



# 产品功能

KuAI的产品使用大致分为几个部分：模型开发与训练、模型管理、模型服务。

模型开发与训练有三种模式可选：第一种是可视化开发，第二种是自动学习，第三种是编码开发。可视化开发灵活、高效，其实现过程为：通过可视化拖拉拽的方式配置组件及组件参数，连线形成pipeline实验并运行实验，实现数据的探索、预处理、特征工程以及模型的训练及评估，从而快速地进行模型开发训练，当实验达到预期效果后，便可保存当前由系统自动生成的模型。自动学习则更加标准、规范，其实现开发过程为：通过标准项目的配置完成模型的自动化构建，核心流程包括选择任务类型和优化目标，以及配置训练和测试数据集，并指定特征列和标签列，选择或使用默认算法，系统将基于配置的逻辑自动化完成模型构建。编码开发自由、灵活，通过手工编码的方式进行数据处理、模型开发和训练。

用户基于开发完成的模型提交模型部署服务后，可对服务进行监控和管理，如在线服务的调用情况、离线服务的调度执行情况、模型测试、服务告警等。

|  |  |
| --- | --- |
| **功能名称** | **功能描述** |
| 可视化建模 | 系统支持组件拖拽连线的图形可视化建模方式，通过选择并连接算法组件，调整组件参数，在画布中构建并运行pipeline实验，实现高效的数据加工以及模型的建设及训练。组件包括数据接入、数据预处理、特征工程、训练、预测、评估，以及金融场景专用组件。 |
| 自动学习 | 系统支持自动机器学习建模，用户指定训练目标，上传并拆分训练和数据集，通过自动化数据处理、特征工程、模型筛选、自动调参，实现机器学习自动化。支持二分类、多分类、回归的自动建模训练。 |
| 编码开发训练 | 系统集成了基于开源的JupyterLab，提供在线的交互式开发调试工具。通过创建开发环境，支持自行编写和调测模型训练代码并基于该代码进行模型的训练。为了方便用户可以查看运行中的开发环境状态，平台提供了开发环境监控功能，对当前环境的资源使用进行实时监控。 |
| 模型部署 | 系统支持部署模型的的离线服务。支持不同周期（日、周、月）的离线服务调度以及超时预警， 并对服务版本进行管理，同时支持查看服务的运行日志、进行补数和模型更新等。  系统支持部署模型的在线服务。支持对实例、资源、日志进行监控，支持服务的版本管理，支持接口的在线测试。 |
| 模型监控 | 模型部署上线后，通常需要对模型的稳定性和准确性进行监控。系统支持对模型服务的稳定性和准确性进行监控。  稳定性支持对预测数据量、PSI、模型结果分布进行监控。准确性支持对KS、AUC、F1等主要指标以及曲线进行监控。 |
| 我的云盘 | 系统提供安全、可靠的对象存储服务。支持将文本、图片、音频等文件上传至云盘，对云盘中的文件夹及文件进行新建、复制、移动、重命名等。 |
| 我的镜像 | 当在建模过程对算法框架、开发库有特殊需求时，预置的算法引擎已经不能满足需求，系统支持使用KuAI自定义镜像功能来制作镜像。 |
| 权限管理 | 系统支持权限管理，按账户进行数据隔离。 |

# 产品优势

## 智能化开发，大幅提升机器学习建模效率

* 一站式：支持丰富的数据探索、预处理、特征工程、模型训练、模型预测、模型评估等组件，可一键快速发布模型服务
* 建模智能：支持自动特征工程和智能调参，根据优化目标自动选择最优模型
* 经验直用：内置丰富的金融场景模型和最佳实践，用户无需费时费力开发，拖拽即可用

## 可解释模型，符合金融机构合规运营要求

* 特征可解释：支持特征的重要性筛选和数据的图形化展现，发现重要影响因素和数据特点
* 模型可解释：支持基于机器学习方法快速生成高解释性的模型

## 高度适应性，满足不同业务场景的使用需求

* 支持单机/分布式，支持直连数据源或对接大数据平台，支持CPU/GPU训练

# 应用场景

* 风控：申请评分卡、风险定价、行为评分卡、催收评分卡、中介代办、授信额度分配、推广作弊、团伙欺诈……
* 营销：用户的全生命周期运营、客户分群、用户价值预测、客户流失预测、客户召回预测、销量预测、产品推荐、授信额度分配……

# 典型案例

* 客户

某消费金融机构

* 业务背景及需求
* 建模效率低下，难以满足日常风控需求；
* 人工置顶风控策略，效果不佳，逾期率居高不下；
* 专业建模人才招聘难，人力成本过高。
* 产品方案

使用KuAI进行评分卡、违约预测等模型的开发，对模型进行一站式的训练、运行、监控、更新维护等。

* 方案成效

模型输出效率提升：模型开发从月级别降到天级别