

XSight 人工智能分析平台

应对海量数据平台的分布式挖掘分析工具

XSight 简介

逸迅XSight 是一款企业级的人工智能分析平台产品，通过大数据技术和人工智能技术的融合，提供包含数据预处理、常规机器学习算法、深度学习框架、模型的评估以及预测等相关服务的一站式机器学习平台，降低大数据挖掘落地成本，并帮助企业创造并提升业务价值。

逸迅XSight能帮助用户真正的处理大数据挖掘问题，加速了大数据挖掘效率，降低了大数据挖掘的成本。逸迅XSight提供了大量的富有创造性的特性加速企业大数据挖掘落地。



图 1 逸迅人工智能技术栈

产品功能

可视化交互界面

逸迅XSight通过可视化的方法，整个流程都可以通过拖拽式操作包括:数据导入、数据探索与预览、数据预处理、特征工程、算法选择等，并提供了常用的数据统计和分析的图表，供用户能够直观的从图形中发掘数据背后的意义。

逸迅XSight还提供了交互式数据探索工具供用户对数据进行实验性的探索工作，实现数据的预处理过程。通过先进的大数据技术，缩短了数据探索的时间，帮助用户实时的对数据进行探索和实验。

并行化大数据挖掘

大数据时代的特点是数据量规模巨大，传统的单机算法已经无法满足大数据时代的需要，我们需要使用并行算法来处理 and 计算大数据。逸迅XSight是一款基于Apache Spark架构的真正意义上的大数据挖掘平台，是完全基于Hadoop和Spark的并行化的平台，其计算能力受限于整个大数据平台的整体计算能力。

当大数据平台的节点得到扩展时，逸迅XSight的计算能力也随之扩展。其通过Spark提供的高性能内存迭代计算引擎在多个节点上并行挖掘，解决了单机节点无法挖掘海量数据的问题，同时提高了挖掘速度。



图 2 XunSight 并行架

逸迅XSight自主研发了大量并行运行在Spark之上的数据挖掘算法。这些算法覆盖了数据挖掘工作中最常用的算法种类。用户可以非常方便的在逸迅XSight中使用这些算法来创建模型。

在大数据挖掘实践中，使用并行化的数据挖掘算法只是整个挖掘工作的一部分，更多的工作集中在数据预处理方面。在大数据环境下，单机处理海量数据显然是不现实的。逸迅XSight提供了多种数据预处理的并行化算法，帮助用户高效的对数据进行处理。

高效的大大数据挖掘落地平台

逸迅XSight专注于提高企业大数据挖掘的效率。在大数据挖掘实践中，往往70%-80%的时间和精力耗费在数据探索和数据处理上面。因此，提高数据探索和数据处理的效率会大大加速整个大数据挖掘的落地速度。

在数据处理方面，逸迅XSight除了提供大量的并行化数据处理算法之外，还对SQL进行了良好的支持。用户可以通过SQL语句对建模的中间结果进行各种操作，这样，具有SQL技能的用户可以零学习成本的使用SQL对数据进行各种探索和处理。

丰富的挖掘算法选择

逸迅XSight是基于Spark架构的并行化数据挖掘平台。逸迅XSight自主研发了大量的并行化数据挖掘算法，这些算法解决了单机算法不能挖掘大数据的问题，极大的方便了用户在大数据上的数据挖掘。

此外，逸迅XSight除了数据挖掘的算法之外，逸迅还提供了若干数据处理的并行化算法，同样也可以在模型工作流的算子中使用，对数据进行并行化的处理。

算法名称	算法类型
决策树	分类算法
梯度提升决策树	分类算法
随机森林	分类算法
Logistic 回归	分类算法
支持向量机	分类算法
多项式朴素贝叶斯	分类算法
回归树	回归算法
梯度提升回归树	回归算法
回归森林	回归算法

算法名称	算法类型
线性回归	回归算法
岭回归	回归算法
回归	回归算法
均值	聚类算法
主成分分析降维	降维算法
极差归一化	归一化算法
去均值归一化	归一化算法
标准归一化	归一化算法
标准向量生成	向量化算法
哈希向量生成	向量化算法
随机采样	采样算法

图 3 逸迅并行算法

高效的数据处理速度

逸迅XSight专注于提高企业大数据挖掘的效率。在大数据挖掘实践中，往往70%-80%的时间和精力耗费在数据探索和数据处理上面。因此，提高数据探索和数据处理的效率会大大加速整个大数据挖掘的落地速度。

在数据处理方面，逸迅XSight除了提供大量的并行化数据处理算法之外，还对SQL进行了良好的支持。用户可以通过SQL语句对建模的中间结果进行各种操作，这样，具有SQL技能的用户可以零学习成本的使用SQL对数据进行各种探索和处理。

产品优势

√ 技术融合

拥有逸迅大数据基因，融合了大数据与人工智能双方技术

√ 简单性

对底层的分布式算法封装，提供拖拽可视化操作环境，实现“搭积木式”的数据挖掘过程。

√ 丰富算法

包含特征工程、数据预处理、统计分析、机器学习、深度学习框架、预测与评估等一整套的机器学习算法组件。

√ 高效挖掘

帮助用户真正的处理大数据挖掘问题，加速了大数据挖掘效率，降低了大数据挖掘的成本。

√ 支持主流框架

包括Tensorflow、Caffe等主流机器学习框架。

典型应用领域：



公安



金融



工业行业



政府



智慧城市



逸迅科技

国内领先的基于大数据的人工智能解决方案供应商



扫一扫 马上关注