

# 联想制造行业大数据解决方案

## 什么样的问题可以被大数据解决

### 举个例子

- A客户的系统里有他们的几十万用户的大量信息，包括各种静态和动态信息、使用记录。每个用户的实际情况都是千差万别。
- 由于业务复杂，A客户目前向这些用户提供的服务也是有琳琅满目的大量选择，前端人员许多都无法熟记并分别。
- 前端人员与用户通过有限时间的沟通，不能完全理解每个用户的情况，只能凭个人经验做出判断做出推荐。
- 如此大量的数据，不太可能由一个人或一组人，通过人力，全局学习这些数据并找出可信的结果。（这个客户数据量并不算大，但是如果要靠一个人来搞清楚，打个比方，就相当去当于几十所大学的校长并对每个学生如数家珍）

## 大数据项目如何实施

### 业务、需求

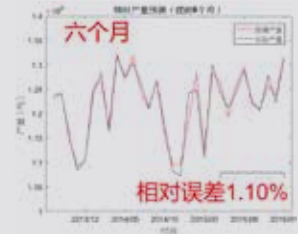
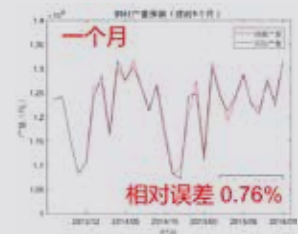
- (大数据) 存储/计算
- (大数据) 查询/报表
- (大数据) 预测/决策

### 智能技术能解决哪些需求

- 数据驱动的
- 可通过数据验证的
- 比下围棋更简单的
- 预测问题

### 客户对数据认识的局限性

### 客户对算法能力认识的局限性



## 联想大数据解决方案



# 联想大数据应用规划

- 大数据方案提供整体系统的架构设计和实施，从梳理业务需求到技术实现，从应用软件到软硬件平台。
- 硬件配置和软件架构都与模型和应用直接相关，至少应满足基本要求。
- 不合理的软硬件配置，尤其是软件平台和工具，如与细分应用无关，则在实施时无法利用。
- 软硬件系统最大化负荷模型需求则可以获得最好的实施效果。

## 业务目标梳理：

- 理解业务需求，梳理建模
- 目标，提供现场规划
- 数据实验，模型原型设计

## 定制软件架构：

- 功能需求分析
- 性能需求分析
- 兼容性分析

## 定制硬件架构：

- 计算界限分析，计算密集型分析
- I/O界限分析，I/O密集度分析数据摄入的要求



# 联想大数据架构设计

## 硬件架构必须考虑到不同的要求：

- 计算界限分析，计算密集型分析  
例如，数据分类，文本挖掘，或先进的运算法则
- I/O界限分析，I/O密集度分析  
例如，索引、搜索、数据排序、数据摄入的要求
- 频繁算法梳理：  
例如，矩阵运算，傅立叶变换

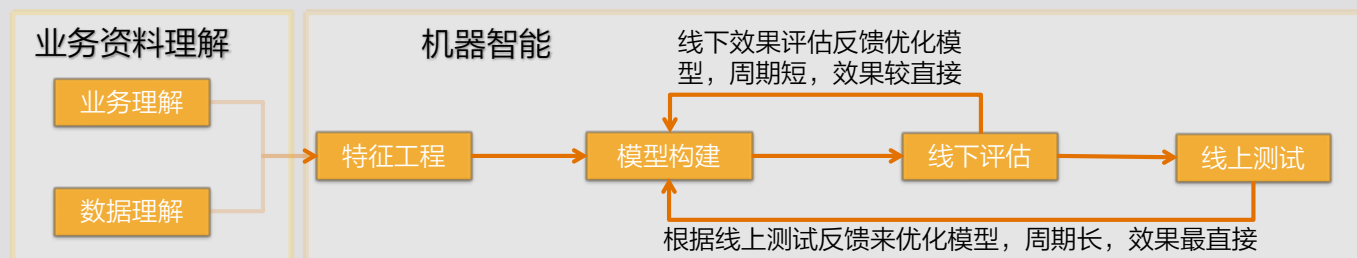


## 基础设计：

- 处理器：时钟速率和核数
- 网络：任务通信开销，时钟时间，多路连接
- 内存：各服务内存开销，频繁任务的计算量
- 存储介质：所选平台软件的SSD支持 硬盘并发数，核对硬盘比

# 联想大数据机器智能实施方法

通过完整的大数据分析，协助企业收集和组合各种相关数据，借助机器学习和知识图谱来发现和探索，发掘出数据和业务的本质联系，支撑企业业务和管理的发展和转型。



## 业务理解

- 理解业务逻辑和商业规则

## 数据理解

- 理解数据字段含义，完整性及清晰度

## 特征工程

- 数据清洗及降噪
- 高维特征生成
- 特征选择及优化

## 模型构建

- 大数据平台存储
- 分布式机器学习模型
- 根据反馈不断优化模型学习能力

## 线下评估

- 设定线下评估机制
- 利用线下评估效果模拟线上真实效果
- 版本迭代成本小，周期短

## 线上测试

- 将结果投放到真实场景中
- 利用真实场景的少量样本进行测试
- 最接近真实效果，但迭代周期长



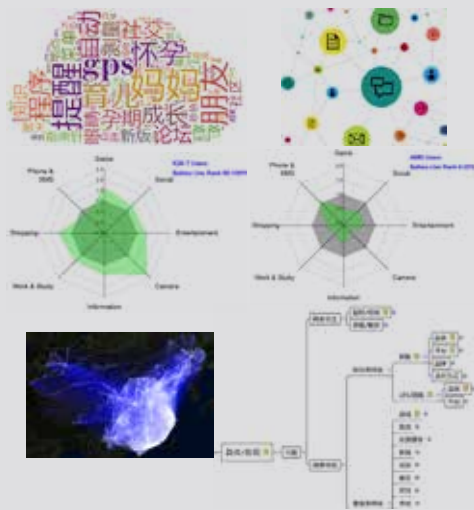
# 联想大数据应用 —— 洞察：用户、应用、设备画像

## 核心算法

• 文本挖掘 • 机器学习 • 知识图谱



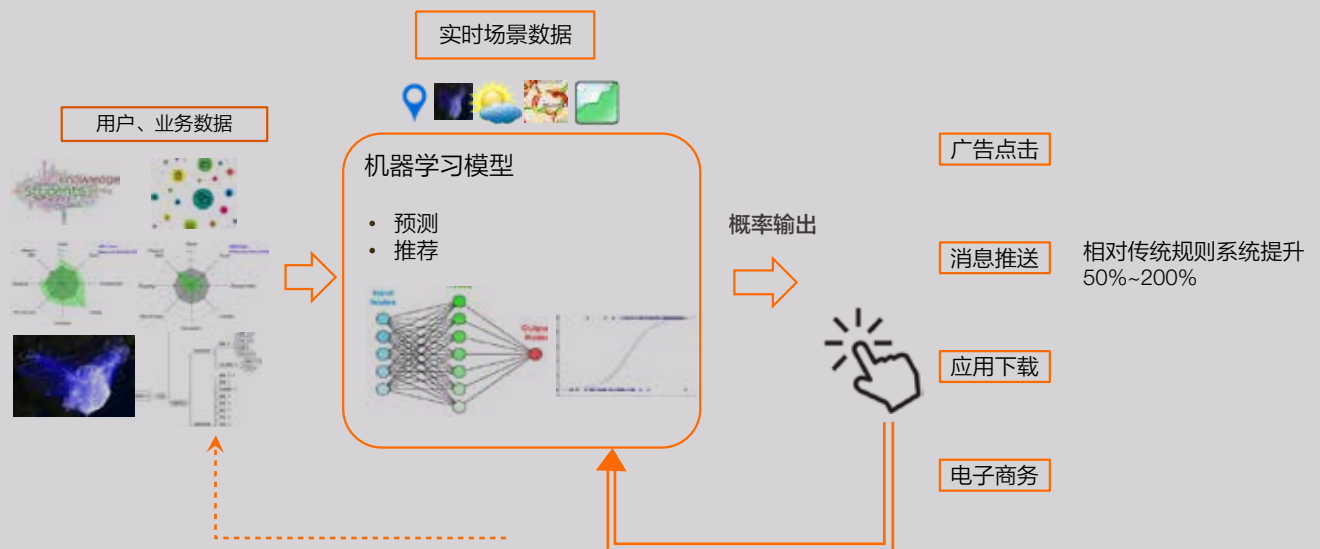
应用画像



设备画像

用户画像

# 联想大数据应用 —— 预测：建立业务闭环



# 联想大数据应用 —— 知识理解：人机问答系统

**Lenovo:** 欢迎使用LENOVO产品自动问答系统。问题示例：  
T450S有什么样的配置  
T450S和T450分别配了什么CPU和显卡？  
Y900能玩使命召唤吗？  
推荐一款16GB内存的电脑吧

**You:** T450S和T450内存分别是多大？  
**Lenovo:** thinkpad t450的memory 可以选择以下配置：8GBx2,8GBx1,4GBx1  
thinkpad t450s的memory (soldered+dimmm) 可以选择以下配置：4+2GB,4+4GB,4+8GB,4+0GB

**You:** 想找个4GB内存的手机  
**Lenovo:** ['thinkpad yoga 11e', 'lenovo yoga 2 pro', 'lenovo yoga tablet 2 with windows 13', 'lenovo yoga 3 14 inch', 'thinkpad yoga 260', 'thinkpad yoga 12', 'thinkpad yoga 15', 'lenovo yoga 3 11 inch', 'thinkpad yoga', 'thinkpad yoga 14', 'thinkpad yoga 11e 2nd gen']

**You:** Y900能玩使命召唤吗？  
**Lenovo:** 我们的资料显示使命召唤 12的推荐配置是i5-4670K CPU,8GB内存。 GTX 970显卡  
ideacentre y900 可能有多种配置可选  
i7-6700K (4C, 4.0 / 4.2GHz, 8MB, 2133MHz) CPU 满足要求  
NVIDIA GeForce GTX970, 4GB memory和NVIDIA GeForce GTX980, 4GB memory GPU 满足要求  
16和32GB 内存满足要求

**You:** ideapad 300有什么颜色。15寸的  
**Lenovo:** ideapad 300 15 inch的top color 可以选择以下配置: Black,Silver,Red

- 知识图谱
- 语义分析
- 知识表示
- 深度学习