

# 深度商业分析在金融行业的实践

**CDA** 数据分析师  
www.cda.cn

鲍忠铁



**TalkingData**

移动 · 数据 · 价值

# 目录

一 深度商业分析的七大基石

二 客户洞察和数据应用

三 机器学习在金融行业应用实践

CDA 数据分析师  
www.cda.cn

# 一 深度商业分析的七大基石

## 从Target预测少女的案例谈起



- Target分析师 安德鲁.波尔 基于网上商店行为，利用数据挖掘技术创建了“怀孕预测指标”
- 波尔的计算机抓去25种商品进行分析，对每个购物者进行打分，还可以预测预产期，根据怀孕所处的各个阶段派送优惠券
- 不断测试，从数据中发现规律，怀孕中期，孕妇会大量购买免香型沐浴露，怀孕20周会大量购买镁钙锌等保健品
- 如果一个客户在3月购买了，可可油沐浴露，大量镁钙锌保健品，淡蓝色小地毯。其可能的分娩日期是8月底。

# 一 深度商业分析的七大基石

## 深度商业分析目的

职能	过去	现在	未来
分析团队	发生了什么	什么正在发生，为何发生？	将来会发生什么
执行团队	该如何做？	正在如何做？	应该如何做？

## 何为深度商业分析？



利用数据和工具找到隐藏的规律，对即将发生的商业行为进行预测，为企业经营提供决策支持。

# 一 深度商业分析的七大基石

## 1 商业挑战基石

- 如何获取新客户和保持存量客户
- 如何定位潜在客户和提升市场份额
- 如何应对外部竞争威胁
- 哪些是企业高利润客户，如何获取更多的高利润客户
- 哪些是企业高忠诚度客户，如何获取更多的高忠诚度客户
- 企业的产品体验和客户体验如何，竞争对手如何
- 客户需要的哪些新产品和服务

商业挑战基石将聚焦企业最重要的需求，引导商业分析战略或方案的发展

# 一 深度商业分析的七大基石

## 2 数据基石

- 将企业内部数据源进行存储和集中（包括商业分析无关的数据，例如调查数据，行业数据）
- 基于商业场景和用户经营，将数据进行标签化，另外可以考虑建立统一的用户管理系统
- 引入外部数据进行数据整合，包括运营商数据、社交数据、移动App数据等
- 建立基于客户的综合视图，微观分析客户兴趣和消费偏好
- 运营指标的标准化、量化、可视化、实时化、趋势化，指导业务运营工作
- 考虑建立数据模型，为决策数据层面的支持

数据是商业分析的基础，需要逐步渐进，逐步提升，内部数据为主，外部数据为辅

# 一 深度商业分析的七大基石

## 3 分析实施基石

- 商业目标是什么，何时达到商业目标
- 达到商业目标的过程中，商业战略是什么
- 执行计划和行动中，需要哪些部门的支持
- 需要哪些工具和软件
- 需要哪些数据作为支撑



以上问题是分析实施前需要考虑的问题，如果处理好，将会搭建成功商业分析基础

# 一 深度商业分析的七大基石

## 4 洞察发现基石

- 训练分析人员超越个体
- 从商业挑战、目标、问题入手
- 只列出所能回答商业问题答案的事实
- 需要了解业务需求和数据来源
- 聚焦数据所含意义而不是数据
- 提供商业建议，无论其多么困难
- 广泛传播、商业效果监测、量化评估

CDA 数据分析师  
www.cda.cn 传播观察



洞察发现的主要目的是发现规律，提供决策支持，传播反馈很重要



# 一 深度商业分析的七大基石

## 5 执行和测算基石

- 描述没有方案之前发生的情况
- 描述方案之后发生的情况
- 方案如何增加商业收益（客户保持、客户获取、CLTV，客户流失、商业收入）
- 方案如何提升企业生产力（例如人均销售、平均订单数、人均外呼数量、营销成本、客户体验）
- 内部客户对方案的满意度（满意度、背靠背评分）

执行和测算主要目标是展示分析工作，是否完成完成目标，确保商业分析不断体现其积极的商业效果。

# 一 深度商业分析的七大基石

## 6 知识共享基石

- 利用企业的智慧从数据中创造更多的商业价值，在企业中培养数据文化，各部门员工主动了解商业分析情报。
- 将企业从命令管理模式转向智慧数据共享模式

## 7 创新基石

深度商业分析的逻辑是

- 揭示
- 预测
- 优化

成功企业的创新特征

- 聚焦客户
- 产品差异化
- 理解并预测市场及趋势
- 执行迅速
- 鼓励员工推行分析和创新

竞争激烈的商业环境，利用暴增的数据信息，分析和创新成为成功的生命线

# 二 客户洞察和数据应用

## 用户画像的原则

- 1 信用信息和人口属性信息为主**    **2 采用强相关信息，忽略弱相关信息**    **3 将定量的信息归类为定性的信息**

信用信息是描述一个人在社会中的消费能力信息，所有的信息都是信用信息就是这个道理。定位金融企业需要触达客户，人口属性信息就是起到触达客户的作用。

强相关信息就是同场景需求直接相关的信息，其可以是因果信息，也可以是相关程度很高的信息

如果定义采用0到1作为相关系数取值范围的化，0.6以上的相关系数就应该定义为强相关信息。

用户画像的目的是为产品筛选出目标客户，定量的信息不利于对客户进行筛选，需要将定量信息转化为定性信息，通过信息类别来筛选人群。



## 二 客户洞察和数据应用

用户画像数据类型分类，共计五类

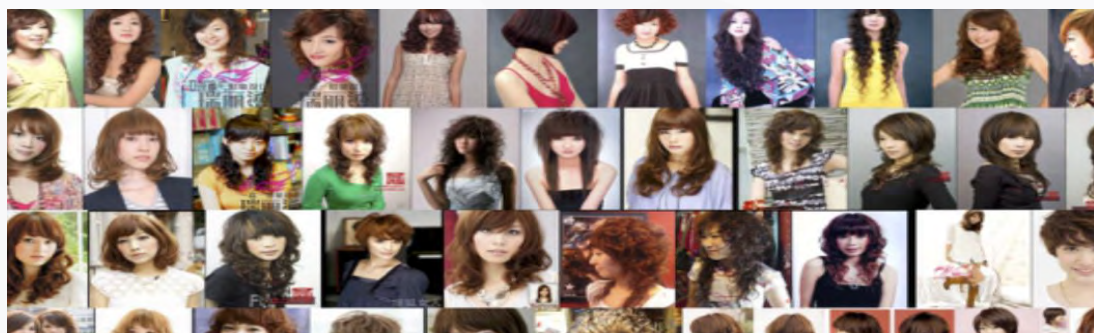
- 1 人口属性——用户是谁，如何联系用户
- 2 信用信息——风险控制（资产，收入，支付习惯）
- 3 消费特征——消费习惯和消费偏好
- 4 兴趣爱好——品牌偏好，个人兴趣

以上所有的信息皆为静态信息，大多数短时间不变，为历史信息，更新频率可以定为一年

- 5 社交属性——社交媒体，言论特点，个人爱好

动态信息，需要及时更新

用户画像涉及数据的纬度需要业务场景结合，既要简单干练又要和业务强相关，既要筛选便捷又要方便进一步操作。



## 二 客户洞察和数据应用

金融企业用户画像的基本步骤如下

1. 画像相关数据的整理和集中
2. 找到同业务场景强相关数据
3. 对数据进行分类和标签化
4. 依据业务需求引入外部数据
5. 按照业务需求进行筛选客户

CDA 数据分析师  
www.cda.cn



用户画像工作刚刚开始，需要外部数据和支持帮你走向光明

## 二 客户洞察和数据应用

# DMP



DMP可以通过技术手段整理内部数据，并将外部数据引入到金融企业内部，建立标准的标签体系，提供灵活的用户画像方式，深层次分析客户，按照业务场景筛选客户，最终触达客户。



**CDA** 数据分析师  
www.cda.cn



## 二 客户洞察和数据应用

### 银行DMP业务数据场景

寻找贷款，消费，财富管理的客户。消费金融和财富管理是重点。

理财用户：定期将工资 / 收入转到支付宝 + 理财App高度活跃用户

购房用户：收入 + 年龄 + 无房产 + 出现在房产销售区域 + 学区房区域

消费贷款人群：年龄 + 收入 + 工资账户余额 + 电商App活跃 + 理财App不活跃 + 贷款App活跃

境外游客户：境外消费 / 机场位置 + 旅游App高活跃用户 + 境外游社交数据

购车客户：年龄 + 收入 + 无车 + 汽车类App高活跃用户 + 汽车论坛言论 + 4S店

分期客户：年龄 + 收入 + 工资 + 月还款金额 + 还款时间

# 三 机器学习在金融行业应用实践

## 聚类分析的应用

### 定义：

聚类分析是把对象划分到不同的类，使得类内的对象尽可能相似，而类间对象尽可能相异。

### 与传统分类相比：

- **科学性：**传统分类依靠经验，对于研究对象大多靠猜；聚类分析依靠客观数据以及其中隐藏的信息把相似的个体聚为同一类
- **全面性：**聚类在分析时考虑对象所有维度上的信息（成百上千），而传统分类一次只能研究几个维度上的信息
- **探索性：**传统分类需要事先定义好划分的数量和特点，聚类分析研究的划分是未知的，需要在研究的过程中不断总结和发现划分的特点





# 三 机器学习在金融行业应用实践

## 聚类分析的应用

### 人群II 僵尸型用户



- 以女性为主，年龄偏小
- 重钱包、轻理财
- 低频使用钱大掌柜应用
- 最后一次登录日期低于第一次日期，可能投资后没有后续登录行为

### 人群III 投资型用户

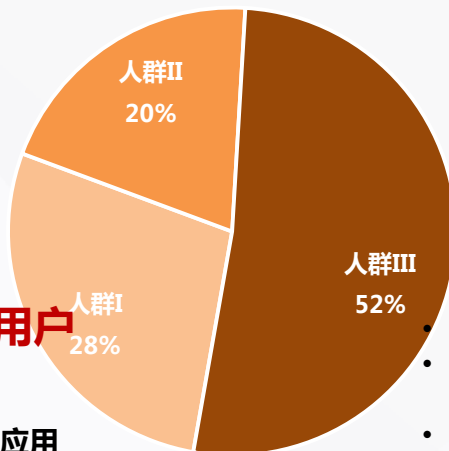


- 年龄偏大，男女适中
- 习惯使用理财、钱包产品，且投资收益大，持仓量高
- 部分用户使用应用在早上10点前
- 登录和使用钱大掌柜频繁

### 人群I 一次性型用户

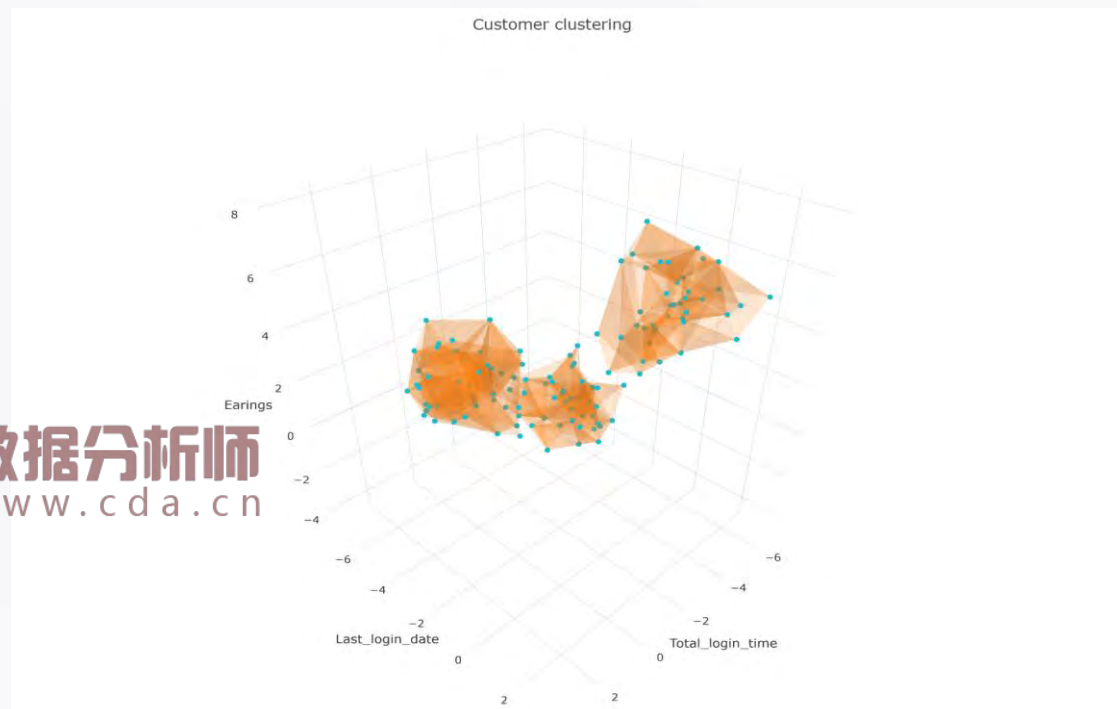


- 年龄适中，男女适中
- 几乎不使用钱大掌柜应用
- 第一次和最后一次时间相近，估计为受营销活动所致，一次性用户



- ❖ 左表为3类人群分类情况，其中数值表示该类此类人群平均值（标准化），数值大小含义请见变量描述
- ❖ 数据已标准化，取之范围在 - 1到1之间，后续对实际数据进行解读

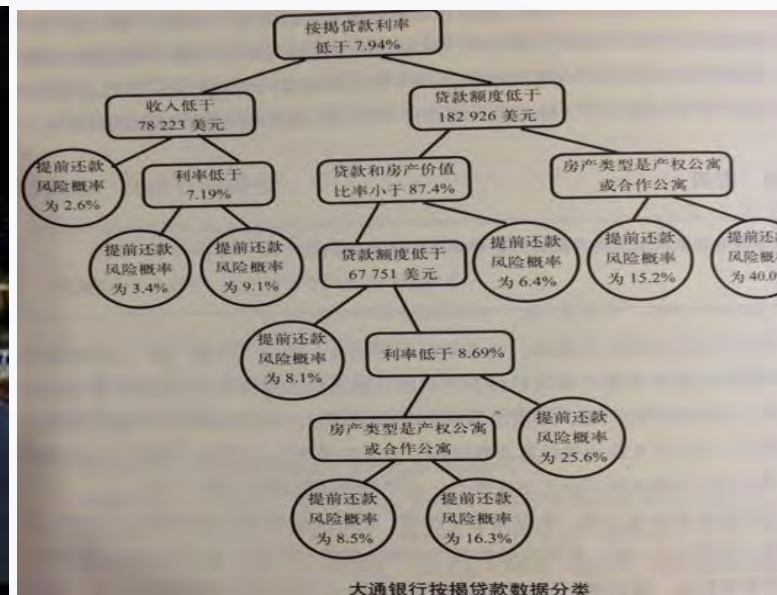
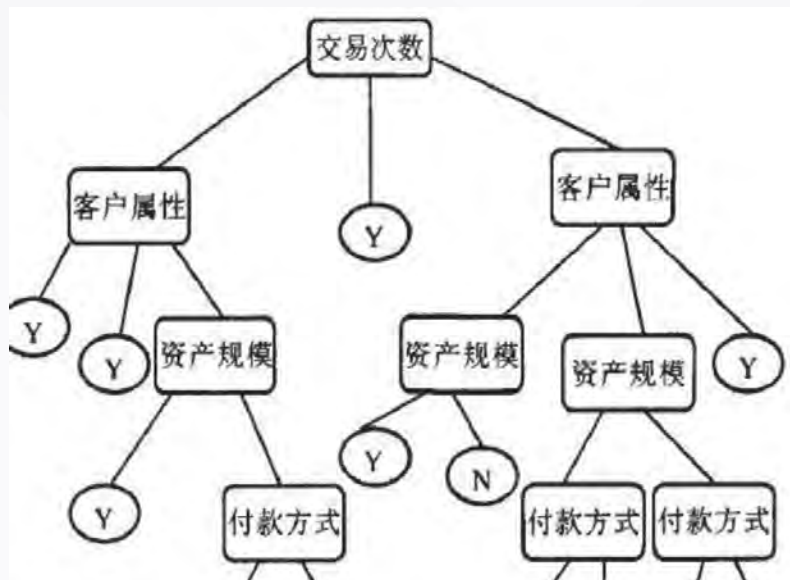
CDA 数据分析师  
www.cda.cn



- ❖ 在3D图像中，x轴代表客户最后一次登陆日期，y轴代表总登陆时间，z轴代表客户收益。
- ❖ X, Y, Z均为抽取的局部变量
- ❖ 黄色的图像标明了根据客户的行为和收益等将客户聚成几类
- ❖ 由图可以看出，用户被分成**3类人群**，和上述结论相符

# 三 机器学习在金融行业应用实践

## 决策树用于抵押贷款管理



- 利用决策树技术，进行按揭贷款风险识别。计算未来客户逾期还款或提前还款发生的概率
- 银行实施提前介入方法，降低逾期还款的概率或将资金提前进行安排，提高资金利用率。项目实施一年来银行增加了6亿美金收入

# 三 机器学习在金融行业应用实践

## Lookalike 算法的应用



高净值客户种子筛选



十亿活跃设备机器学习



潜在用户定向广告投放

短信转化率 6.5%， 人均购买理财产品12万， 千元营销费用带来亿元理财产品购买

高净值客户客挖掘和营销

# 三 机器学习在金融行业应用实践

## 基于IBM沃森的财富产品推荐



应用

- 利用IBM 沃森机器计算结果，为高端客户提供适合其风险偏好的财富产品

收益

- 减低了营销成本
- 提高了客户满意度
- 提高客户的收益
- 吸引了更多的高端客户

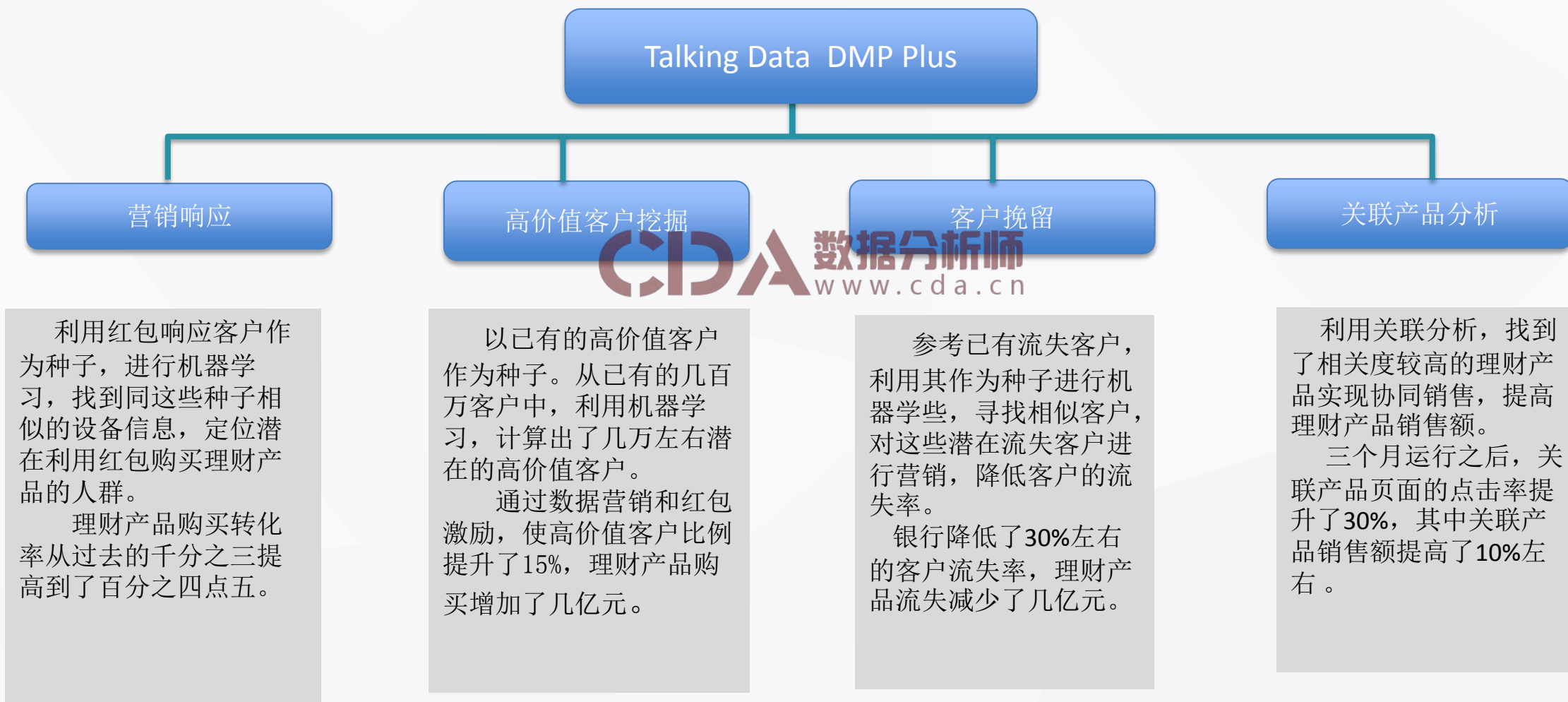
数据

- 银行自身数据
- 客户提供的数据
- 外部互联网数据
- 各项经济数据

CDA 数据分析师  
www.cda.cn

# 三 机器学习在金融行业应用实践

## 机器学习在银行客户经营上的应用



# 三 机器学习在金融行业应用实践



- 1 伦敦证券交易所 40%的交易是由算法和模型驱动的
- 2 美国的John Elder（约翰，埃尔德）博士建立了股票预测的黑盒子，押上了全部身家—10年
- 3 美国三角洲金融公司开发了交易系统，预测标准普尔指数变动，自称准确率为70%—骗局

# 三 机器学习在金融行业应用实践

## 看驾驶习惯买车险



UBI就是Usage Based Insurance，基于驾驶行为而定保费的保险，保费取决于驾驶时间、地点、驾驶方式等综合指标考量，为记录驾驶员的上述行为并关联理赔金额，参加UBI的车主都会在车上安装UBI车载智能盒子。欧洲和北美以及韩国早已颁布了一些鼓励政策以促进UBI发展，中国也已于2012年核准发布了UBI产品。

CDA 数据分析师  
www.cda.cn

OBD是英文On-Board Diagnostic的缩写，中文翻译为“车载诊断系统”。这个系统随时监控发动机的运行状况和尾气后处理系统的工作状态，一旦发现有可能引起排放超标的情况，会马上发出警示。

美国好事达保险公司，通过2012的建模大赛，实现了仅仅通过保险车辆特征将人身伤害赔偿责任预测精确度提高3倍，每年可以给公司节约4000万美元成本。

01

## 金融行业

招商银行 中国银行 兴业银行 浦发银行 平安 泰康保险 国泰君安 华泰证券 国信证券 ...

02

## 房地产行业

碧桂园地产 龙湖地产 卓越地产 北京城棚

03

## 零售行业

欧时力 老百姓大药房

## 品牌客户



TalkingData

移动 · 数据 · 价值

04

欧莱雅 宝洁 惠氏 达能 宝马 北汽

## 航空酒店

05

国航 东航 深航 春秋 华住

06

## 互联网金融

宜信 玖富 翼龙贷 国美金控 万达金融

07

## 开发者市场

聚美优品, 今日头条, 墨迹天气等8万款应用  
消灭星星, 全民枪战, 花千骨等2万款游戏

08

## 国际市场

摩根斯坦利, 老虎, 红杉资本等40+顶级投行VC,  
Google/Facebook/Twitter/Yahoo移动广告归因对接

CDPA 谢谢!  
数据分析师  
www.cda.cn