



江南一号大数据分析 与建模系统建设蓝图

(版本 3.0)

KDS Global, LLC
2018 年 8 月 8 日

目录

一、引言.....	2
二、背景介绍.....	2
三、KDS 公司介绍以及行业优势.....	3
(一) 公司介绍.....	3
(二) 行业优势.....	4
四、江南一号大数据分析建模系统.....	13
(一) 江南一号硬件软件以及数据需求.....	14
(二) 江南一号各部门大数据整合系统.....	15
(三) 江南农商行吸纳及扩大支付宝芝麻客户群的风险与收益的管理系统.....	18
(四) 江南一号个人 ABC 风险收益定价智投系统.....	21
(五) 江南一号 ABC 资管智投系统.....	23
(六) 账户信用评分及价值评分模型.....	25
(七) 小额贷款模型.....	29
附录一： 江南一号系统建模所需清单.....	32
附录二： 江南一号大数据分析与建模系统定价标准.....	37

一、引言

此文旨在介绍 KDS Global, LLC (以下简称 KDS 公司) 为江南农商行大数据分析建模系统建设的蓝图设计。将 KDS 公司在国际金融市场 30 余年的先进的大数据智能决策的技术和经验加以应用, 为江南农商行建立一套大数据分析建模系统, 达到防范风险、提高精细化管理和盈利能力的目的。

文章主要包括以下几个内容: 背景介绍、KDS 公司介绍和技术优势、以及江南一号大数据系统分析及建模系统的介绍。

二、背景介绍

世界已经进入大数据时代。大数据渗透到了每一个行业和业务领域, 成为重要的生产因素。数据每天都在高速膨胀, 数以亿记的变化。可以说谁掌握了大数据, 谁就占领了市场的制高点, 赢得了先机, 就能立于不败之地。

金融业是全世界各个行业中最依赖于数据的, 而且最容易实现数据变现的行业。当下, 探索智能化转型已成为国际银行业的共识。如何提高大数据的掌握和运用能力, 已经成为金融机构未来发展的核心竞争要素。随着巴塞尔协议 III 的颁布, 近年来央行对金融监管力度不断加大; 自 2015 年利率市场化后, 利差下降也使得靠传统利差收益变的艰难, 整个银行业的经济利润下降, 对产品创新和提升综合服务的需求日益迫切。金融监管的要求和市场变化的需求是

银行业加大风险管理力度，提高精细化管理水平，增强盈利能力的机遇和挑战。

在这样的时代背景下，建立一套银行自有的基于大数据的智能决策系统非常必要。此建设蓝图旨在将 KDS 世界领先的大数据智能决策核心科技在江南农商行进行复制、创新和运用，建立江南一号大数据分析建模智能决策系统，提升资源配置效率，强化风险管控能力，促进业务创新，提高盈利水平。同时也为将来建立整个农商行的网络系统数据库奠定基础。帮助各农商行提升大数据分析和运用能力，以达到利润最大化。

三、KDS 公司介绍以及行业优势

(一) 公司介绍

KDS 的经营理念：敏行如思，顺逆合一，紧握概率，循道蓄德。

KDS 是一家硅谷金融科技公司，拥有世界一流的大数据、人工智能、云计算专利技术平台，高性能的大数据挖掘，系列数学模型（包括数据挖掘、量子场论、微分几何以及流形拓扑）、区块体技术以及多项金融模型专利，为全球顶尖的金融机构提供金融科技服务。

在过去的 30 余年里，KDS 为近百家世界知名银行提供优质的金融服务，并为他们带来了高额的收益回报。这些客户使用我们的智能决策系统管理的总资产超过数万亿美元。我们的客户包含商业银行（富国银行、美国银行、摩根大通、花旗银行等）；中央银行和监管机构（美国联邦储备局、中国外汇管理局）；投资银行（摩根

斯坦利、高盛等) 以及机构性与非机构性投资者 (Annaly、黑石等)。

(二) 行业优势

1. 丰富的国际金融市场大数据应用经验

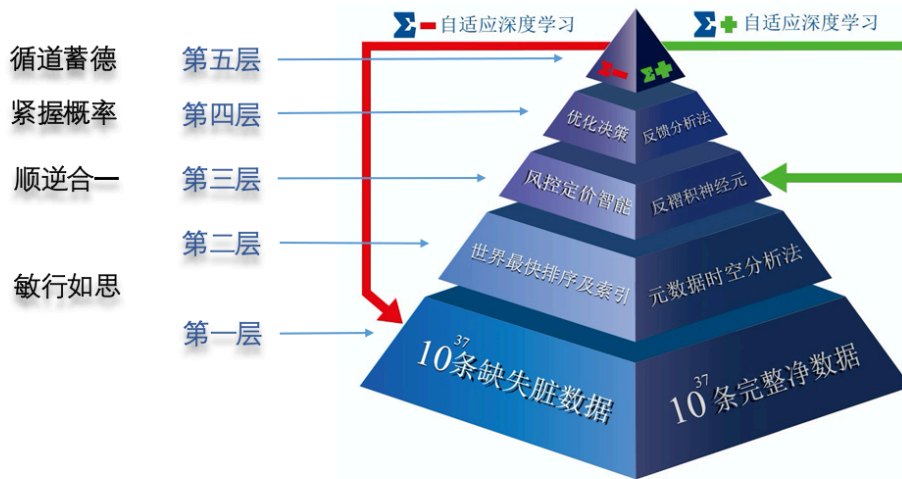
在过去的 30 余年里, KDS 的大数据分析以及建模的能力在美国华尔街的近百家知名银行中得到检验。美国富国银行是美国个人住房贷款发放额度最大的银行, KDS 在长达十余年的时间里, 作为富国银行的合作伙伴, 为其上万亿的贷款基金做风险分析。每月 KDS 基于实时更新的智能决策系统, 为富国银行实时生成近百份不同类型的风险分析以及定价报告。这些多维度的深度报告包括个人贷款提前还款及违约风险分析、不同银行贷款风险分析、新生贷款以及历史交易风险分析等等。随着 KDS 神经网络映射技术新时代的到来, ABC UBXTM 黎曼流形几何式区块体 (看不见的密码) 可以作为一个 N 维欧几里得几何式区块面 (可看见的密码) 运作, 而不是以传统线性链的形式, 这项技术可以支持银行总行多方同时访问分布式分类账, 增加信息安全性, 大力提高银行结算和清算效率, 以及增强预测未来盈利的能力。

2. 智能决策一体化的 ABC UBXTM 全系统

KDS 的 ABC UBXTM 系统是一个基于五层金字塔的全系统。ABC 分别代表: A (AI Network Engine) 人工智能网络发动机, B (10^{37} Big Data) 大数据, C (Cloud Computing) 云计算。UBX 代表的是

无处不在，无时不在的对大数据每一个数据在系统内均排序并全局索引，达到敏行如思、顺逆合一、紧握概率、循道蓄德的多项世界专利的思想和技术能力。

ABC UBXTM 人工智能发动机网络



第一层：原始数据的采集、清洗、提炼、转换、导入ABC UBXTM系统

第二层：建立元数据库、ABC UBXTM专利排序索引及分析—敏行如思

第三层：人工智能反褶积神经元的风险定价模型建立—顺逆合一

第四层：优化经营交易决策实时反馈分析—紧握概率

第五层：自适应深度学习来优化模型—循道蓄德

目前，多数公司大数据应用技术专注于其中某一层或两层，例如美国数据公司1010公司专注的是第一二层的数据库使用，EMBS公司主要在建立第三层的模型，Bloomberg彭博则是在第四层的交易决策。KDS的核心ABC UBXTM系统是一个利用人工智能、大数据和

云计算实现从原始数据的采集、清理、提炼、转换、加载导入系统、排序索引（KDS专利）、建立模型、目标管理、智能决策、反馈分析等全功能一体化的全系统。这个系统包括了基础数据的收集到顶层的决策，是个全方位的系统工程。我们的大数据系统更先进系统化。

3. 快速实时分析

KDS 的核心智能决策系统 ABC UBXTM 拥有世界先进的排序索引技术，和传统大数据系统相比，数据的处理运行速度有着绝对优势。当进行海量数据计算分析时，可以瞬时产生风险分析报告，真正的做到了敏行如思。公司基于实时金融大数据制作的风险分析报告多年来居于华尔街首位（45 秒内产生报告发放市场）。第二位发布数据的公司比 KDS 的时间晚达 5 倍的时间。在处理更大数据例如 10^{37} 的关系中，ABC UBXTM 排序引擎可以比行业内速度快达 100 倍的时间。这项优势在线上小额贷款审批、客户风险实时评分以及实时动态查询时均会发挥功效。

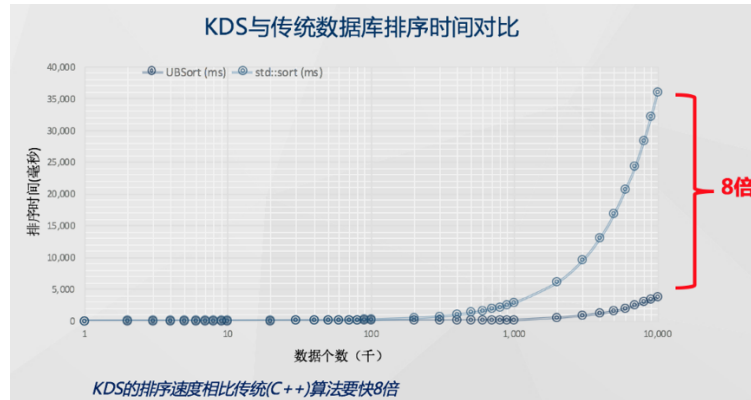
下页是关于 KDS 的核心 ABC UBXTM 与传统大数据系统的优势对比：

ABC UBXTM 系统和传统数据排序性能对比

- UBSort 是我们的专利排序算法（专利号# 5,278,987）
- 时间复杂度为 $O(N)$
- 传统的经典排序算法时间复杂度为 $O(N * \log(N))$
- 数据量越大，UBSort 算法的表现越好

- 性能测试对比了 UBSort 和 C++标准库中的排序模块
- 测试机器为 Intel(R) Xeon(R) CPU E5-2637 v2 @ 3.50GHz, 16 核, 512GB 内存
- 操作系统为 CentOS 6.7, 使用 gcc-4.x, libstdc++.so.6.0.13
- 测试内容为对 1000 万个双精度浮点数进行排序的时间

ABC UBXTM性能案例1：KDS与传统数据库排序时间对比 一敏行如思



在从第一层“原数据”到第二层“元数据”的转换中，ABC UBXTM系统数据的高速全局排序索引比传统C++算法呈指数级别的差别。上图展示的是在英特尔或英伟达芯片下，KDS比传统数据库的排序速度快8倍。可重构矩阵计算机下的ABC UBXTM系统，将比传统数据库有着更为显著的优势。下图就可重构矩阵计算机下的ABC UBXTM系统速度与传统系统的速度做了具体的对比。

纪录数	10^6	10^7	10^8	10^9	10^{10}	10^{11}	10^{12}	10^{1055}
UBXTM比传统系统快的倍数	20	23	27	30	33	37	40	3500

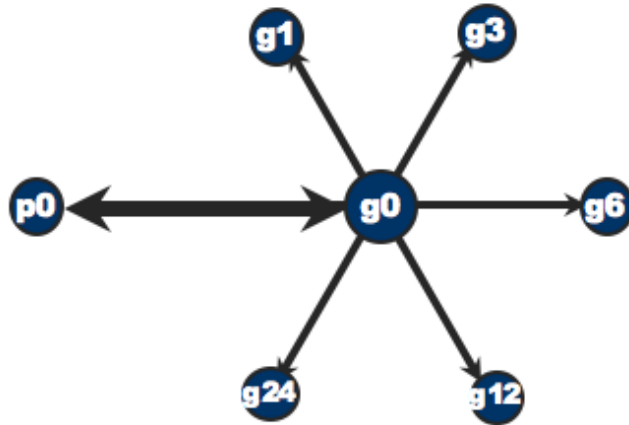
举例来说：A股市场有3500支股票，那么所有的投资组合数目是 2^{3500} 个，亦即 10^{1055} 种投资组合。在对这些关系进行排序时，可重构矩阵计算机下的ABC UBXTM系统的排序算法比传统数据库排序算法快3500倍。

针对目前初步采集到的江南农商行的60T的数据仓库，ABC

UBX™将比传统的速度快2400倍（见上表所示 10^{12} 即1T快40倍，60T将快2400倍）。换句话说，传统数据库下7天才能完成的数据计算，ABC UBX™只需4分钟就可以完成；或是传统数据库下三个月需要完成的模型调试，在可重构矩阵计算机下的ABC UBX™系统下只需要一个小时就可完成，真正做到了敏行如思。高速的排序能力是顺逆合一和紧握概率的必要条件。不难试想，在几个小时内完成的大数据运算和模型校准，比三个月运行一次的模型校准更能精准实时地顺逆合一做到紧握概率，这在闪电式小额贷中的交易决定会有很大的应用空间。

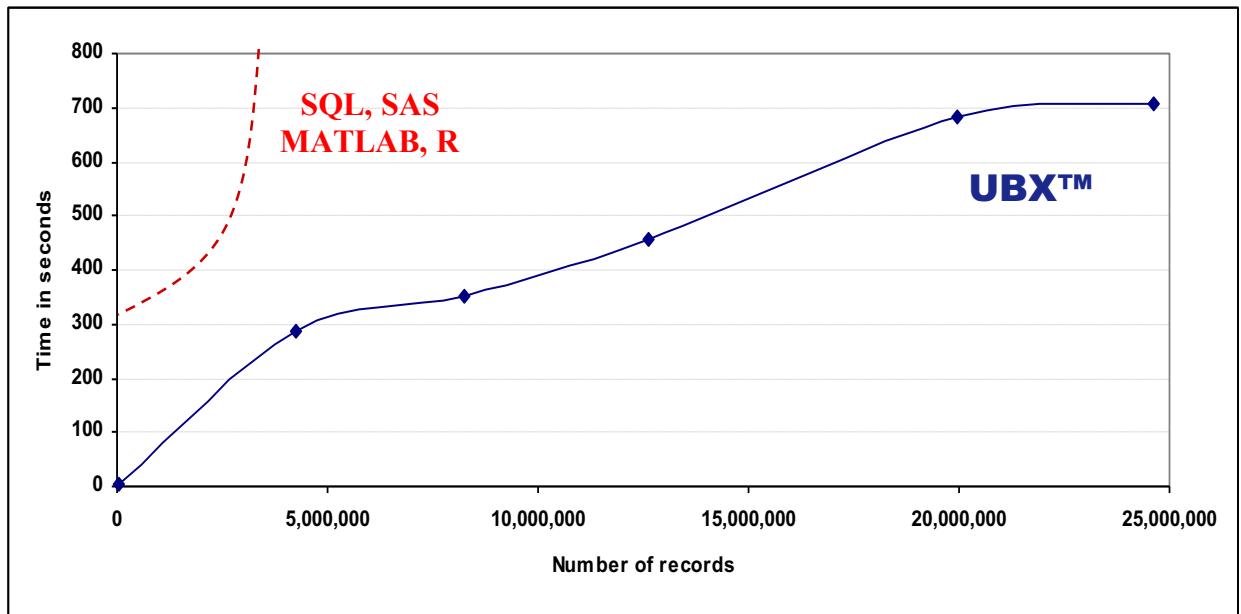
ABC UBX™性能案例 2：星型连接的速度对比 — 敏行如思

- 3000 万条记录，3.2GB 数据库
- 多连接键的相等连接和左外连接
- 多个累积变量
- 超过 3 亿次计算



SQL	86小时	4 CPU
SAS	6小时	4 CPU
UBX™	4分钟	4 CPU

ABC UBXTM 性能案例 3：非线性最小二乘回归的速度对比—紧握概率



记录条数	数据大小 (MB)	数据文件个数	分布式节点数	NLIN 迭代次数	所用秒数
45,889	3.15	1	6	9	5
4,254,142	896.61	12	6	8	288
8,243,801	1,737.48	24	6	8	353
12,606,708	2,657.02	36	6	8	456
19,953,262	4,205.39	60	6	8	682
24,621,612	5,189.30	83	6	8	709

4. 准确性

KDS 有着丰富的金融产品建模的经验，模型的准确性在 30 余年来历经了华尔街市场的检验。KDS 基于机器学习和反褶积神经元技术的风险评分以及定价模型，不同于流行的反映历史数据的风险评分，创新的结合了信贷的宏观经济因子、贷款概况、服务机构信息以及贷款标准的相互作用带来的市场风险。模型评分每日更新，一个分数就可以全面了解未来提前还款或违约风险状况。KDS 精准的模型，不仅建立在先进的建模技术、丰富的建模经验的基础上，同时也是得益于 ABC UBX™ 系统的完整与高速。单就 KDS 风险评分模型来说，数据纵跨了超过 25 年的历史年份，查询了约有 100 亿行的数据，34 列特征字段，14 个子模型。ABC UBX™ 第一层和第二层数据处理的完整性以及高速的数据处理能力，为第三层的模型准确性建立提供了高质量的数据基础，节省并赢得了模型建立、调试、优化的时间。

5. 全维度分析易操作

ABC UBX™ 实现各部门之间数据串联集中化管理，客户对数据可以做任意层面切割分析，客户画像全面灵活。可以做历史时间序列分析、时效曲线分析、客户开户静态起始点分析、流失客户分析、利差分析等等。客户使用可视化页面进行数据查询以及动态实时报告生生，完全不需写代码。

四、江南一号大数据分析及建模系统

在设计江南一号大数据分析及建模系统的时候，需要强调的是：大数据的精髓不是随机抽样，而是全体的数据。不是因果关系，而是相关关系。江南一号在大数据的使用上，存在两个重要节点，一是最底层基础数据的处理，包括整合、清洗、导入以及数据统一平台的建设，即ABC UBXTM系统的第一层和第二层。二是挖掘数据背后隐藏的规律、相关关系，建立模型。因此，第一层和第二层大数据数据库的建立乃是建模以及制定经营决策的基础，这两层数据的清理以及完整程度直接影响到模型的精准程度。在前面的ABC UBXTM系统金字塔中，我们也看到当第五层也就是最上层的经营决策反馈分析的结果与模型预期不符的时候，系统会回到第一层的原始数据重新开始排查原因。因此，基础数据的完整在模型建立中至关重要。

根据 ABC UBXTM人工智能发动机网络图（见第 5 页），江南一号大数据分析及建模系统的工作将分为以下几个步骤完成：

- 建立系统的硬件软件环境
- 数据准备、清洗、转换、导入系统——第一层 敏行如思
- 将导入系统的数据用 ABC UBXTM 排序索引、完成数据库的建立以及前台可视化用户界面的建立——第二层 敏行如思
- 江南农商行吸纳及扩大支付宝芝麻客户群的风险与收益的管理系统、江南一号个人 ABC 风险收益定价智投系统、江

南一号 ABC 资管智投系统、账户信用评分以及价值评分模型、小额大款模型以及更多模型的建立——第三层 顺逆合一

- 经营决策、交易决策以达到风险管控、产品创新提高经济利润——第四层 紧握概率
- 决策反馈分析，根据财务报告实时跟踪模型的结果，相应调整模型——第五层 循道蓄德

这里将着重从下面七个内容描述：

- 江南一号硬件软件以及数据需求
- 江南一号各部门大数据整合系统
- 江南农商行吸纳及扩大支付宝芝麻客户群的风险与收益的管理系统
- 江南一号个人 ABC 风险收益定价智投系统
- 江南一号 ABC 资管智投系统
- 账户信用评分以及价值评分模型
- 小额贷款模型

(一) 江南一号硬件软件以及数据需求

建立江南一号，所需的硬件要求：

2 个主控服务器和 8 个分布式服务器，共 10 个服务器，具体配置要求如下：

2 个主控服务器，每个需要：

CPU: 2*12 核 Intel(R) Xeon(R) E5 26 系

RAM: 512GB DDR4 ECC

HD: 8*2TB SATA 6GB/s

8 个分布式服务器，每个需要：

CPU: 2*4 核 Intel(R) Xeon(R) E3 12 系

RAM: 256GB DDR4 ECC

HD: 2*2TB SATA 6GB/s, Raid 1

操作系统：Centos 7 Linux 发行版最新稳定版

数据要求请见附录。

(二) 江南一号各部门大数据整合系统

江南一号大数据整合系统将整合串联江南农商行各部门之间的数据，建立起一套各部门及总行共用的大数据系统，构建起一套客户、营销及财务等不同场景的 360 度立体画像。

系统环境建立以及数据采集、清洗、转化并导入后，将使用 ABC UBX™ 全系统排序索引完成数据库建立以及完成前台可视化用户界面。这个时候，江南一号大数据系统的第一层和第二层的工作已经完成，用户即可以开始进行实时的动态的数据切割、全维度挖掘分析。下面是江南一号各部门大数据整合系统的部分应用描述：

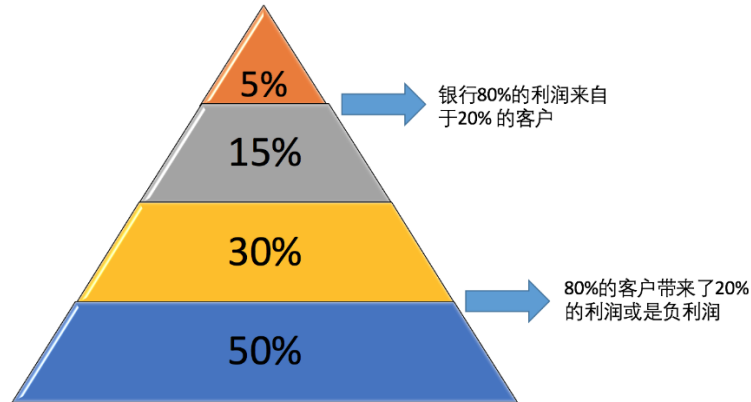
1. 满足监管风险管控需求。运用现有大数据，对不同情境下的信用风险、市场风险、操作风险、流动性风险以及商业风险进行压力测试，以达到巴塞尔协议 III 以及央行监管的的相关要求。KDS

在美国金融市场有丰富的为知名商业银行服务 CCAR 检验的经验（CCAR: comprehensive capital analysis and review 全面资本分析与检验，是美联储最重要的年度压力风险测试，测试银行在极度恶劣的市场环境中是否有足够的资本维持运转）。

2. 内部资金转移价格（FTP）的应用。作为内部考核和管理资产负债的一个有效工具，内部资金转移价格能够科学的识别资产负债的利润来源，可以帮助银行实现利率定价和利率风险管理。江南一号大数据系统可以基于已有数据计算配置内部资金转移定价系统的基础数据，建立内部资金转移价格模型。如果银行已经建立此模型，江南一号将整合 FTP 数据并拓展数据间的关系计算作为 ABC UBX™ 第一层和第二层的输入，做更深次的挖掘，完善优化已有的 FTP 模型。

3. 精细化管理。在江南一号大数据系统下，价值管理将落实到每一个客户，甚至客户下的每一个产品的价值创造情况。颗粒度越小越能协助银行进行优质客户筛选、客户选择、定价管理以及资本配置等精细化管理。每个客户的所有行内交易足迹（开户、存款、贷款、理财、信用卡、还款、消费等交易记录）将与第三方数据整合，利用 KDS 评分系统为每一个客户进行风险及价值打分，根据对银行经济利润贡献度不同分组，针对不同组的客户群，制定不同的策略提供针对化的服务。基于客户的综合贡献，例如持有产品数、净收入贡献、未来的潜力等，对客户进行分类，明确战略类、潜力

类客户；建立涵盖客户不同业务的综合贡献视图，例如客户过去的综合贡献度信息，最后依据客户的需求引导战略客户发展多业务、多产品合作关系，巩固客户的忠诚度，持续提高对银行的贡献度。



一些应用的描述：

(1) 每个客户盈亏业绩报表、信用风险评分和消费习惯分类

(2) 对客户分类（按利润贡献率、年龄组，职业，收入组，价值评分等），分析消费习惯聚类、客户坏账率、金额坏账率、客户循环率、每笔提现金额、每户每月消费次数、年化率总利息、年化率总收费、年化率手续费、年化率净收益、不同时间段的平均消费额、平均提现额、平均贷款额等等。制定相应战略，提供差异化服务，将上图金字塔中 30%的客户群转换为 15%，15%的客户群转换为 5%的优质客户。

(3) 客户流失率分析（什么时候流失的，流失到哪里去了，流失客户有无理财产品等）历史流失率流失情景分析，预测客户流失可能性：估算客户的风险价值。

(4) 客户留存率分析。分析客户留存策略的有效性、估算客户的风险价值。

(5) 客户价格敏感度、新开账户虚假欺诈可能性分析。

(6) 违约可能性严重性分析、交易欺诈可能性分析达到优化违约管理。

(7) 服务渠道分析、交叉销售分析。对客户在不同平台上（柜台、网上、手机、电话营销、微信、线上）的交易足迹分析，可为银行创造更多的价值挖掘。使用多个渠道的客户更多可能拥有银行的多种产品，对这些客户可以提供交叉产品业务，提高客户留存率。

(8) 分析方式：时间序列分析（单个或分类客户过往的交易足迹）、利差分析、客户开户时状态分析（静态分析）、逐笔账户或集合分析等。

（三）江南农商行吸纳及扩大支付宝芝麻客户群的风险与收益的管理系统

在市场经济的大潮中，有个著名的“蝴蝶效应”理论，即初始条件的十分微小的变化，对其未来状态往往可能会造成极其巨大的差别和连锁反应。在客户的信用模型的建立中，每一个维度数据的增加与缺失，都可能对模型的结果产生指数级的影响。因此，数据采集的完整性、准确性，以及实时的更新数据，对模型的准确和经营决策的制定至关重要。

江南农商行通过和第三方征信系统芝麻评分，取得了较好的成效，目前江南农商行拥有近六百万的客户，芝麻信用下拥有约四千万客户，随着业务的继续合作，仍有着巨大的市场扩展空间和业务挖掘的潜力。

芝麻信用评分系统基于以下五个定性的维度计算出客户的一个综合信用分数：行为偏好、信用历史、人脉关系、履约能力和身份特征。这五个维度代表了 10^5 的定性、非定量的非线性关系。芝麻分数每月更新一次。

随着江南农商行和芝麻信用的合作，市场会不断的扩展，客户会不断增加，仅仅借助于第三方的风险评分，不确定因素很多，风险也较大。因此，建立一套自有的风险管控模型，科学的产品定价机制，有效的计算风险调控收益非常重要。

KDS 建立江南农商行吸纳及扩大支付宝芝麻客户群的风险与收益的管理系统将比第三方的评分系统更为先进、精准全面、数据更新更及时，运营操作更安全。其优势具体展现在：

- 1) **模型评分更科学更为精准。**芝麻评分是基于 5 个定性的维度，而我们的模型包括至少 10 个定量的维度，即可以达到 10^{10} 到 10^{37} 的非线性定量的关系（客户基本信息、消费习惯、资金渠道、贷款原因、贷款方式、贷款利差、季节性、宏观经济利率差等等），将能捕捉到更多更深层次的数据关系，从而能使模型更客观更精确。芝麻信

用的五大定性维度输出将作为江南一号信用模型中的输入，加上江南一号各部门打通后建立的大数据系统中的客户的所有特征信息，以及其他外部数据利用 ABC UBXTM 三位一体的统一场论，用机器语言深度学习、验证、复制并优化，大大提高了模型的精确性。

- 2) **模型评分更加全面。**芝麻分数为单个客户计算出一个综合风险信用分数，KDS 的模型会对每个客户计算出四个不同维度的信用风险分数：提前还款的风险、拖欠还款的风险、违约风险和坏账风险。并根据市场利率价格用反褶积神经网络每天实时更新定价，更全面、精细的量化描述了未来客户主动提前还款或是拖欠、违约给银行带来损失的概率，为精准定价提供了保障。
- 3) **评分日更新更及时反映变化。**相比于芝麻分数每个月为单个客户计算出一个综合分数，KDS 的每天实时更新每个客户的四个风险分数：提前还款的风险、拖欠还款的风险、违约风险和坏账风险，日更新可以实时的反映市场利率价格的变化，为江南农商行评估市场风险、资本风险以及利率风险下的资本定价提供了数据的支持和科学的依据。

- 4) **自有模型更安全。**模型完全搭建在江南农商行的内部系统，对数据的安全性和保密性都提供了保障。建模过程对江南农商行完全开放，不会有“黑箱操作”。

这个系统还可为将来拓展借呗花呗乃至还呗的网贷业务提供技术支持。信贷客户未来还款的现金流可以使用信贷资产证券化的方式来运营管理，达到进一步控制风险和提高银行盈利的目的。KDS 在资产证券化的领域有着丰富的国际金融市场的先进经验。

建模所需时间： 3—6 个月

(四) 江南一号个人 ABC 风险收益定价智投系统

利率的市场化对依赖利息收入为主要营收来源的银行业带来了巨大的冲击，银行的收益受到了严重的影响。银行业亟需建立一套科学的产品定价水平、提高客户的精细化管理、产品创新和增加中收能力，以弥补利差缩小带来的影响。

目前银行业的定价水平还处于初步发展的阶段，在五大行及全国性的股份制银行中不乏出现下面的情况：风险资产增多而净利息收益却下降；问题资产增多而资产减值却减少；对高风险行业定价甚至低于低风险行业；银行经济利润排名与资产规模排名不符等。这些问题反映了定价管理的不完善或失效。江南一号个人 ABC 风险收益定价智能系统将基于江南一号的全系统的平台化优势，为江南农商行提供定价管理，包括：内部移转定价 (Fund Transfer Pricing)、风险定价能力、存款的定价能力、资产组合的定价能

力、中间业务的定价能力、客户的定价能力，以及加强内部定价检核的管理。

江南一号个人ABC风险收益定价智能系统将建立在江南一号大数据系统之上，基于ABC UBXTM系统第一层和第二层的大数据对所有客户进行全面画像的个性化精细化管理。2017年，中国大陆有约14亿人口，其中最顶层的2.8亿人的平均收入是最低层2.8亿人的11倍。顶层客户群之外的11亿人是巨大的市场，江南一号ABC风险收益定价智投系统为社会各阶层客户提供个性化的资产组合方案，真正做到普惠金融，运用人工智能、 10^{37} 大数据的高速实时处理能力以及云计算的强大的投研体系和模型策略，利用马克维茨有效边界模型、夏普比率、反褶积神经网络系统等，不论任何市场情况下（顺逆合一）实时的从股市和基金里自动择选出风险调节后的最佳收益的排列组合方案。这个为客户量身打造最优化的资产配置方案根据客户的价值评分、风险承受能力、投资期限、收益预期、单项以及组合产品定价，包括现金类管理、固定收益类投资、权益类投资以及其他智能匹配的金融产品。

ABC UBXTM的 10^{37} 维度的大数据的深层挖掘，以及实时交易决策追踪系统可以实时监控到变化，进而及时调整模型。江南一号个人ABC风险收益定价智投系统基于江南一号全系统的优势，为每个单一客户提供更精确的、实时跟踪风险收益，根据银行日结算的周期相应更新评分，并能实时做出相应调整的、基于目标投资组合的资

产管理配置策略，从而能更实时的反应风险的变化达到精准决策的目的。

建模所需时间： 3—6 个月

(五) 江南一号 ABC 资管智投系统

在 7x24 全球交易市场环境中，哪家银行如果能够及时、准确、并且自动得获取有关过去、现在和未来现金流增值、风险调整价格以及市场询价与投标交易价格的预期变化的大数据，无疑就能够在同业中赢得先机和领跑市场。建立在 ABC UBXTM 系统上，KDS 的现场可编程门阵列（FPGA）的物联网可重构矩阵引擎（RME）事物中以流水线方式实施马克维茨投资组合优化、DNN 和反褶积算法，以及区块体技术安全记录、保存、汇总总账并且根据实时数据完成复杂的实时风险回报生成，是国际领先的资产配置解决方案。

江南一号 ABC 资管智投系统是 KDS 根据江南农商行的收益预期和风险偏好，结合中国 A 股市场的实际情况，利用 ABC UBXTM 人工智能系统为江南农商行量身定做的一套基于 A 股基金产品的资管智投系统。

在过去 30 年中，在统一量子场论（QED）的指导下，通过在国际金融市场不断积累与实践，KDS 在 ABC UBXTM 人工智能系统中建立并发展了一系列强有力的金融建模方法。从功能角度看，可以划分为收益模型、风险模型和优化模型；从交易频率看，可以划分为高频模型和低频模型。ABC UBXTM 高频收益模型可以对股票或基金

在下一时刻的价格（间隔不大于 1 分钟）进行预测；ABC UBXTM 低频收益模型可以对股票或基金未来一个月的收益进行预测；ABC UBXTM 风险模型可以对股票或基金在未来一段时间的波动进行预测。UBXTM 优化模型综合考虑收益和 风险，确定最优的投资组合，得益于 KDS 独有的反褶积算法，UBXTM 优化模型可以处理百万级别的数据。

江南一号 ABC 资管智投系统可以做到：

- 1) **高收益。** 随着交易数据积累和交易结果的滚动反馈，ABC UBXTM 人工智能系统对 A 股的分析 and 预测会越来越好，从而在保证预期收益率的前提下，收益率会逐步提高，达到甚至超过 36%。
- 2) **严格风险控制。** 模型的最终输出结果是一个最优投资组合集合，集合中的每个元素是一个投资组合。每个组合都对应着相应的风险，且是在这个风险下收益最高的组合。江南农村商业银行可以根据自己的预期收益和风险偏好，选择适合自己的投资组合。
- 3) **交易自动化。** 得益于 ABC UBXTM 人工智能系统强大的计算能力，每个交易日都可以为下一个交易日计算出最优投资组合。一个简单的交易程序就可以实现自动化交易。自动化交易可以降低基金的运营成本还可以避免人为操作的失误。

KDS 会在整个建模过程中对江南农村商业银行完全开放，防止“黑箱操作”的发生。

建模所需时间：3-6 个月

(六) 账户信用评级及价值评分模型

1. 产品的输出：

对于江南农村商业银行的每个账户单独给出评分，其中贷款账户给信用评分，存款及理财账户给价值评分，对于既是贷款账户又是存款及理财账户给两种评分。引入两种评分体系的原因是，对于贷款账户，因为有借贷行为，风险主要为信用风险，分析现有及历史数据的侧重点在于还款表现上，并以此为依据对未来每个账户的还款风险进行预测，从而给出信用评分。而对于存款和理财账户，其行为的本质是投资，即银行需要向开户人或企业支付利息。这些账户由于没有借贷行为，信用评分不再适用，分析的侧重点在于现有及历史账户为银行创收多少和银行的存款利率及理财产品与市面上别家银行和机构的竞争对比，从而预防重要的、有价值的客户流失的风险。评分的初级输出以表格形式存在，即账户号+评分存放在表格中。每月结算时对新增账户评分对已有账户更新评分，如果按天结算评分可以每天更新。

2. 产品的输入：所需数据清单

(1) 贷款账户数据（即所有有贷款行为的账户，包括已经还贷完毕的历史账户）：

a. 个人账户信息 -----（例如但不限于）账户编号、开户人性别、年龄、工作证明、收入证明、资产证明、信用评级、债务占收入比例等等

b. 企业账户信息-----（例如但不限于）账户编号、企业类型、所属行业、规模、员工人数、市值、资产负债等等

c. 贷款信息 -----（例如但不限于）贷款类型、贷款原因、贷款渠道、有无抵押物、抵押物类型及价值、贷款额、贷款利率是浮动还是固定、初始贷款利率、贷款起始时间、总贷款期限、是否联合贷款、贷款人数等等

d. 每月还贷信息 -----（例如但不限于）每月应该还款额、每月实际还款额、当期贷款利率、贷款剩余额、贷款已还期数、贷款剩余期数、是否逾期未还款、连续几期未还款等等

(2) 存款账户数据（即所有有存款行为的账户，包括已经销户的历史账户）：

a. 个人账户信息 -----（例如但不限于）账户编号、开户人性别、年龄、工作单位、职位、证件类型等等

b. 企业账户信息-----（例如但不限于）账户编号、企业类型、所属行业、规模、员工人数、市值、资产负债等等

c. 存款信息-----（例如但不限于）存款类型、存款时期、年利率、存款期限、计息方式、剩余期限等等

d. 现金流 -----（例如但不限于）存取款日期、存取款金

额、存取款行为发生后的账户剩余金额、销户日期、销户原因等等

e. 创收信息（即如何使用这笔存款来创造收益以支付存款利息）
-----（例如但不限于）银行使用这笔存款创收的渠道、方式、
回报率等等

（3）理财账户数据（即所有投资理财账户，包括已经销户的历史账户）：

a. 个人账户信息 -----（例如但不限于）账户编号、开户人性别、年龄、工作单位、职位、证件类型、风险偏好、投资预期收益等等

b. 企业账户信息-----（例如但不限于）账户编号、企业类型、所属行业、规模、员工人数、市值、资产负债、风险偏好、投资预期收益等等

c. 投资理财产品信息-----（例如但不限于）发行规模、认购规模、发行时间、产品类型、年化收益率、期限、利率、利息是固定或浮动、付息方式、风险等级等等

d. 每期付息信息 -----（例如但不限于）剩余期限、应计天数、应计利息、到期收益率、销户日期、销户原因等等

e. 创收信息（即如何使用这笔投资金来创造收益以支付理财利息）
-----（例如但不限于）银行使用这笔投资金创收的渠道、
方式、回报率等等

（4）市面上其他银行和金融机构的理财产品信息（包括现有

的和历史的):

(例如但不限于) 发行银行或机构、发行规模、认购规模、发行时间、产品类型、年化收益率、投资期限、利率、利息是固定或浮动、付息频率、风险等级等等

(5) 宏观经济数据 (包括当前及历史):

a. 房屋价格指数、成屋销售指数、居民收入指数、居民消费指数等等

b. Shibor 利率、利率互换利率

c. 远期利率

3. 建模计划时间安排

第一阶段 (一个月): 数据准备阶段, 主要工作为

查看、分析、筛选江南农村商业银行所有账户历史数据, 找出建模需要的账户数据

获取相关宏观经济数据 (包括各种利率、房屋价格指数、消费指数等等), 生成远期利率曲线

第二阶段 (三周): 建模阶段, 主要工作为

提取建模所需输入数据

估计模型参数

第三阶段 (五周): 为所有账户给出信用评分和价值评分阶段, 主要工作为

完成宏观经济数据模型 (为打分模型的一个组成部分) 的代码

完成打分模型的代码

为所有账户给出评分

第四阶段（一周）：评分结果测试和验收阶段

（七）小额贷款模型

1. 模型输出

小额贷款模型是一个白名单推荐模型，分为两个部分：

（1）对每个江南农村商业银行存量金融资产个人账户且在本行没有贷款给出小额贷款额度。

（2）对每个江南农村商业银行的信用卡客户给出消费贷款额度以及分期期数。

2. 模型输入——建模所需的数据

（1）定期存款账户

个人基本信息：性别、年龄、学历、婚姻状况、工作证明、收入证明、资产证明、信用评级、债务占收入比例等等。

存款信息：开户日期、存款期限、利率、存款额度等等。

账户行为：月均存款余额、月均取款余额、定期转活期、账户注销、账户注销原因等等。

（2）活期存款账户

个人基本信息：性别、年龄、学历、婚姻状况、工作证明、收入证明、资产证明、信用评级、债务占收入比例等等。

存款信息：开户日期、利率、存款额度等等。

账户行为：月均余额、月均存款次数、月均存款金额、最大存款金额、月最大存款、月均消费金额、月均取款金额、月均转账次数、月均转账金额、月均最大转账金额、当期月份账户余额等等。

（3）国债账户

个人基本信息：性别、年龄、学历、婚姻状况、工作证明、收入证明、资产证明、信用评级、债务占收入比例等等。

国债信息：国债名称、国债利率、计息方式、国债期限、到期日等等。

账户行为：买入国债时间、卖出国债时间、卖出国债金额等等。

（4）理财产品账户

个人基本信息：账户编号、开户人性别、年龄、工作单位、职位、证件类型、风险偏好、投资预期收益等等。

理财产品信息：发行规模、认购规模、发行时间、产品类型、年化收益率、期限、利率、利息是固定或浮动、付息方式、风险等级等等

付息信息：剩余期限、应计天数、应计利息、到期收益率、销户日期、销户原因等等。

（5）住房抵押贷款账户

个人基本信息：性别、年龄、学历、婚姻状况、工作证明、收入证明、资产证明、信用评级、债务占收入比例等等。

贷款信息：贷款额度、贷款期限、贷款利率、还款方式、每月

还款金额、贷款到期日等等。

账户行为：提前还款次数、提前还款金额、逾期次数、月均还款金额等等。

外部数据：

外部征信数据

芝麻信用

违法犯罪记录

失信被执行

以及其他已有外部相关数据。

3. 建模时间安排

第一阶段（4周）：数据准备阶段，主要工作为查看、分析、筛选江南农村商业银行所有账户历史数据，找出建模需要的账户数据

第二阶段（4周）：建模阶段，主要工作为提取建模所需输入数据

估计模型参数

第三阶段（4周）：对金融资产账户和信用卡账户给出额度为每个存量金融资产个人账户给出个人贷款额度。

为每个信用卡客户给出个人消费贷款额度及分期期数。

第四阶段（2周）：模型结果测试和验收

附录一： 江南一号系统建模所需清单

一、建模数据需求

(一) 评分模型

1. 贷款账户数据（即所有有贷款行为的账户，包括已经还贷完毕的历史账户）：

个人账户信息 -----（例如但不限于）账户编号、开户人性别、年龄、工作证明、收入证明、资产证明、信用评级、债务占收入比例等等

企业账户信息-----（例如但不限于）账户编号、企业类型、所属行业、规模、员工人数、市值、资产负债等等

贷款信息 -----（例如但不限于）贷款类型、贷款原因、贷款渠道、有无抵押物、抵押物类型及价值、贷款额、贷款利率是浮动还是固定、初始贷款利率、贷款起始时间、总贷款期限、是否联合贷款、贷款人数等等

每月还贷信息 -----（例如但不限于）每月应该还款额、每月实际还款额、当期贷款利率、贷款剩余额、贷款已还期数、贷款剩余期数、是否逾期未还款、连续几期未还款等等

2. 存款账户数据（即所有有存款行为的账户，包括已经销户的历史账户）：

个人账户信息 -----（例如但不限于）账户编号、开户人性别、年龄、工作单位、职位、证件类型等等

企业账户信息-----（例如但不限于）账户编号、企业类型、所属行业、规模、员工人数、市值、资产负债等等

存款信息-----（例如但不限于）存款类型、存款时期、年利率、存款期限、计息方式、剩余期限等等

现金流 -----（例如但不限于）存取款日期、存取款金额、存取款行为发生后的账户剩余金额、销户日期、销户原因等等

创收信息（即如何使用这笔存款来创造收益以支付存款利息）-----（例如但不限于）银行使用这笔存款创收的渠道、方式、回报率等等

3. 理财账户数据（即所有投资理财账户，包括已经销户的历史账户）：

个人账户信息 -----（例如但不限于）账户编号、开户人性别、年龄、工作单位、职位、证件类型、风险偏好、投资预期收益等等

企业账户信息-----（例如但不限于）账户编号、企业类型、所属行业、规模、员工人数、市值、资产负债、风险偏好、投资预期收益等等

投资理财产品信息-----（例如但不限于）发行规模、认购规模、发行时间、产品类型、年化收益率、期限、利率、利息是固定或浮动、付息方式、风险等级等等

每期付息信息 -----（例如但不限于）剩余期限、应计天

数、应计利息、到期收益率、销户日期、销户原因等等

创收信息（即如何使用这笔投资金来创造收益以支付理财利息）
-----（例如但不限于）银行使用这笔投资金创收的渠道、方式、回报率等等

4. 市面上其他银行和金融机构的理财产品信息（包括现有的和历史的）：

例如但不限于）发行银行或机构、发行规模、认购规模、发行时间、产品类型、年化收益率、投资期限、利率、利息是固定或浮动、付息频率、风险等级等等

5. 宏观经济数据（包括当前及历史）：

房屋价格指数、成屋销售指数、居民收入指数、居民消费指数等等

Shibor 利率、利率互换利率

远期利率

（二）小额贷款线上审批模型

1. 定期存款账户

个人基本信息：性别、年龄、学历、婚姻状况、工作证明、收入证明、资产证明、信用评级、债务占收入比例等等。

存款信息：开户日期、存款期限、利率、存款额度等等。

账户行为：月均存款余额、月均取款余额、定期转活期、账户注销、账户注销原因等等。

2. 活期存款账户

个人基本信息：性别、年龄、学历、婚姻状况、工作证明、收入证明、资产证明、信用评级、债务占收入比例等等。

存款信息：开户日期、利率、存款额度等等。

账户行为：月均余额、月均存款次数、月均存款金额、最大存款金额、月最大存款、月均消费金额、月均取款金额、月均转账次数、月均转账金额、月均最大转账金额、当期月份账户余额等等。

3. 国债账户

个人基本信息：性别、年龄、学历、婚姻状况、工作证明、收入证明、资产证明、信用评级、债务占收入比例等等。

国债信息：国债名称、国债利率、计息方式、国债期限、到期日等等。

账户行为：买入国债时间、卖出国债时间、卖出国债金额等等。

理财产品账户

个人基本信息：账户编号、开户人性别、年龄、工作单位、职位、证件类型、风险偏好、投资预期收益等等。

理财产品信息：发行规模、认购规模、发行时间、产品类型、年化收益率、期限、利率、利息是固定或浮动、付息方式、风险等级等等

付息信息：剩余期限、应计天数、应计利息、到期收益率、销户日期、销户原因等等。

4. 住房抵押贷款账户

个人基本信息：性别、年龄、学历、婚姻状况、工作证明、收入证明、资产证明、信用评级、债务占收入比例等等。

贷款信息：贷款额度、贷款期限、贷款利率、还款方式、每月还款金额、贷款到期日等等。

账户行为：提前还款次数、提前还款金额、逾期次数、月均还款金额等等。

5. 外部数据

外部征信数据

人行征信数据

芝麻信用

违法犯罪记录

失信被执行

反欺诈数据

由江南信息技术部提供建模所需数据源，具体需要的表和字段，以及其他相关建立大数据系统所需的相关数据（财务数据等）由KDS江南工作组和江南信息技术部一起协商决定，保证数据和整个在建模区的建模过程都在江南提供的安全、可控、有监管的环境中进行。

附录二： 江南一号大数据分析建模系统定价标准

江南一号大数据分析建模系统蓝图包括建立江南一号各部门大数据整合系统、江南农商行吸纳及扩大支付宝芝麻客户群的风险与收益的管理系统、江南一号个人ABC风险收益定价智投系统、江南一号ABC资管智投系统、账户信用评分以及价值评分模型、小额贷款模型。

江南一号大数据分析建模系统定价标准分为固定收费或浮动收费两种形式。其中固定收费是按照年收费等方式收取固定费用，浮动收费是按照江南农商行与KDS协商制定的按照经济利润的增幅比例收取费用（江南一号大数据系统的模型输出会链接账表）。收费项包括下面三项内容：

- ABC UBX™ 专利使用费用
- 建模费用
- 江南一号大数据系统建立以后的日常维护、数据更新以及新的应用开发的费用

下面将逐项说明。

1. ABC UBX™ 专利使用费用

在过去的30余年里，KDS始终秉承持续创新的精神、坚信核心科技是科技企业的发展动力，经营收入中大部分都是用于研发投入，拥有了多项世界领先的高科技专利（列表见下），并以此在华尔街金融市场上为世界知名的银行多年来连

续创造了丰厚的收益。

江南一号大数据系统中 ABC UBXTM 专利使用内容包括：

- KDS 的 ABC UBXTM 系统在 30 年来持有的所有专利，以及现在正在研发中的和未来的专利都将完全向江南农商行开放使用于江南一号大数据系统。
- KDS 的 ABC UBXTM 系统的所有软件的使用权
- 系统建立过程中 KDS 总部的首席科技官直接带队的技术团队全程支持。
- 江南一号大数据各部门整合系统。
- 江南一号系大数据系统搭建完成后的系统调试及培训
工作。

以下是 KDS 部分专利列表：

- UBXTM 虚拟袋线性排序与索引
- QED 期权定价模型
- 统一微分经济学
- 可用 TBA-Only 的 MBS 供应 (ATOMSTM)
- UBiquitous IndeX (UBXTM)
- 反褶积风险评分
- 统一 TSPTM 指数收益和价格算法
- 风险调整收益率 (RAYTM) MBS 定价
- 统一 ATOMSTM 指数收益和价格算法

- CMO 现金流规则库生成器

收费标准：固定收费方式 1, 000, 000 美元 / 年；或浮动收费方式，按照协定的江南农商行在使用江南一号大系统后利润增长的比例收费。

2. 模型费用

此项内容包含下述模型以及未来建立的模型：

- 江南农商行吸纳及扩大支付宝芝麻客户群的风险与收益的管理系统
- 江南一号个人 ABC 风险收益定价智投系统
- 江南一号 ABC 资管智投系统
- 账户信用评分以及价值评分模型
- 小额贷款模型

收费标准：250, 000 美元 / 单个模型；或浮动收费，按照协定的江南农商行在使用模型后利润增长的比例收费。

3. 日常系统维护、数据更新以及新的应用开发的费用

此项内容包括江南一号大数据系统建立以后的如下工作：

- 江南一号大数据系统的日常系统维护
- 数据日更新、月更新等
- 基于江南一号大数据系统新的应用开发
- 系统使用问题解答及人员培训

收费标准：100, 000 美元 / 月