



京东JIMI 

用户未来意图预测 让机器人更智能

京东 人工智能平台研发部 邹波

JIMI (JD Instant Messaging Intelligence) 基于AI与大数据的智能机器人

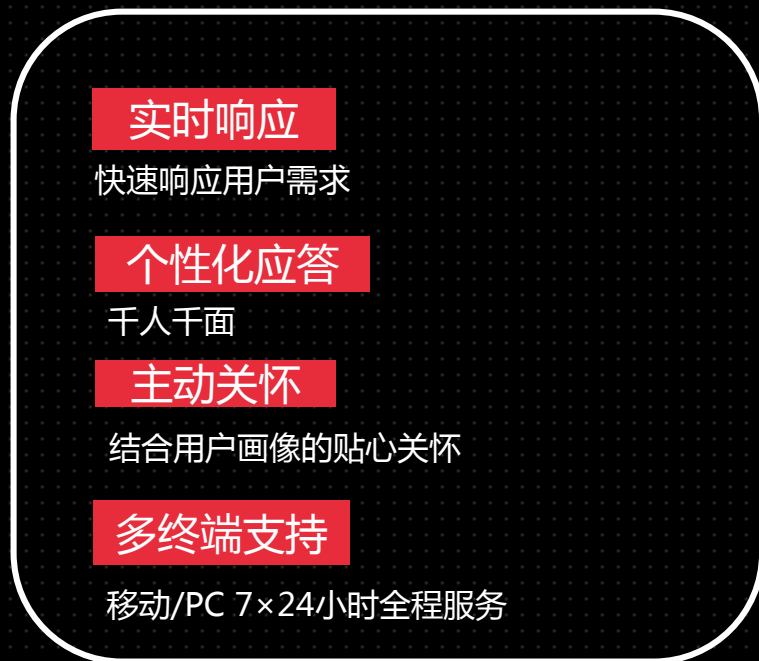
用户需求



大数据

用户数据	订单信息	账户信息	聊天意图	聊天日志	访问记录
其他数据	商品数据	物流数据	评价数据	活动数据	网络数据

智能服务



传统客服助手升级，五大角色助力成长

购物助手

到货提醒

活动咨询

砍价

商品下单

库存咨询

负责提供商城售前咨询服务，带来新式的购物体验

客服助手

配送物流

订单状态

退换货

返修

服务单

安装

价保

负责**分担商城客服的咨询压力和成本**
售后业务种类覆盖率**80%+**

金融助手

借贷

黄金

基金

个人资产

理财推荐

负责为京东用户提供金融类咨询及自助服务

个人助理

笑话

天气

手机充值

百科知识

出行助理

负责提供丰富的百科内容和精彩的热点活动，提升产品趣味

开放平台

模型训练

Badcase
挖掘

数据分析

自动化
测试

语料维护

负责开放技术成果，提供全方位人工智能技术接入方案

专业的核心技术实力

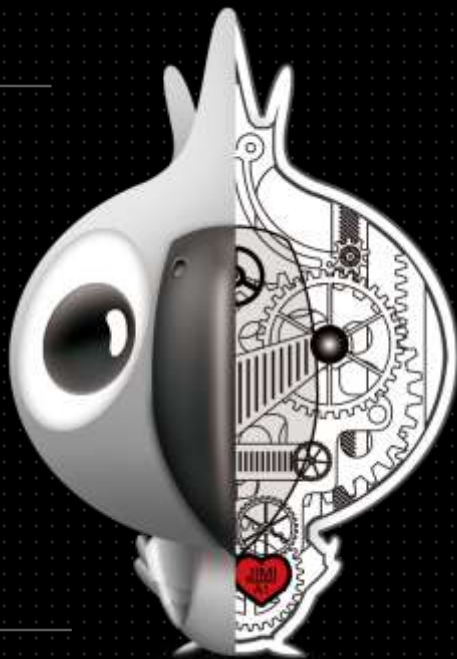
卷积神经网络
Convolutional Neural Network

深度问答系统
DeepQA

异步消息投递
Asynchronous message

一问多答技术
Multiple Answers Technology

抽象消息模板
Abstract Message Template



深度神经网络
Deep Neural Network

机器学习
Machine Learning

知识图谱构建
Mapping Knowledge Domain

用户画像技术
Persona

自然语言处理
Natural Language Processing

多元咨询意图识别

从多种角度理解咨询需求

语义级别问答匹配

答案匹配过程引入语义计算更精准

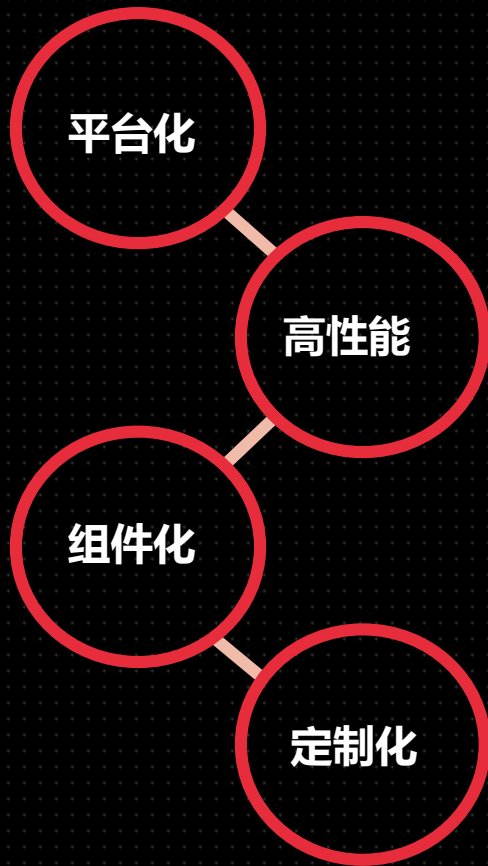
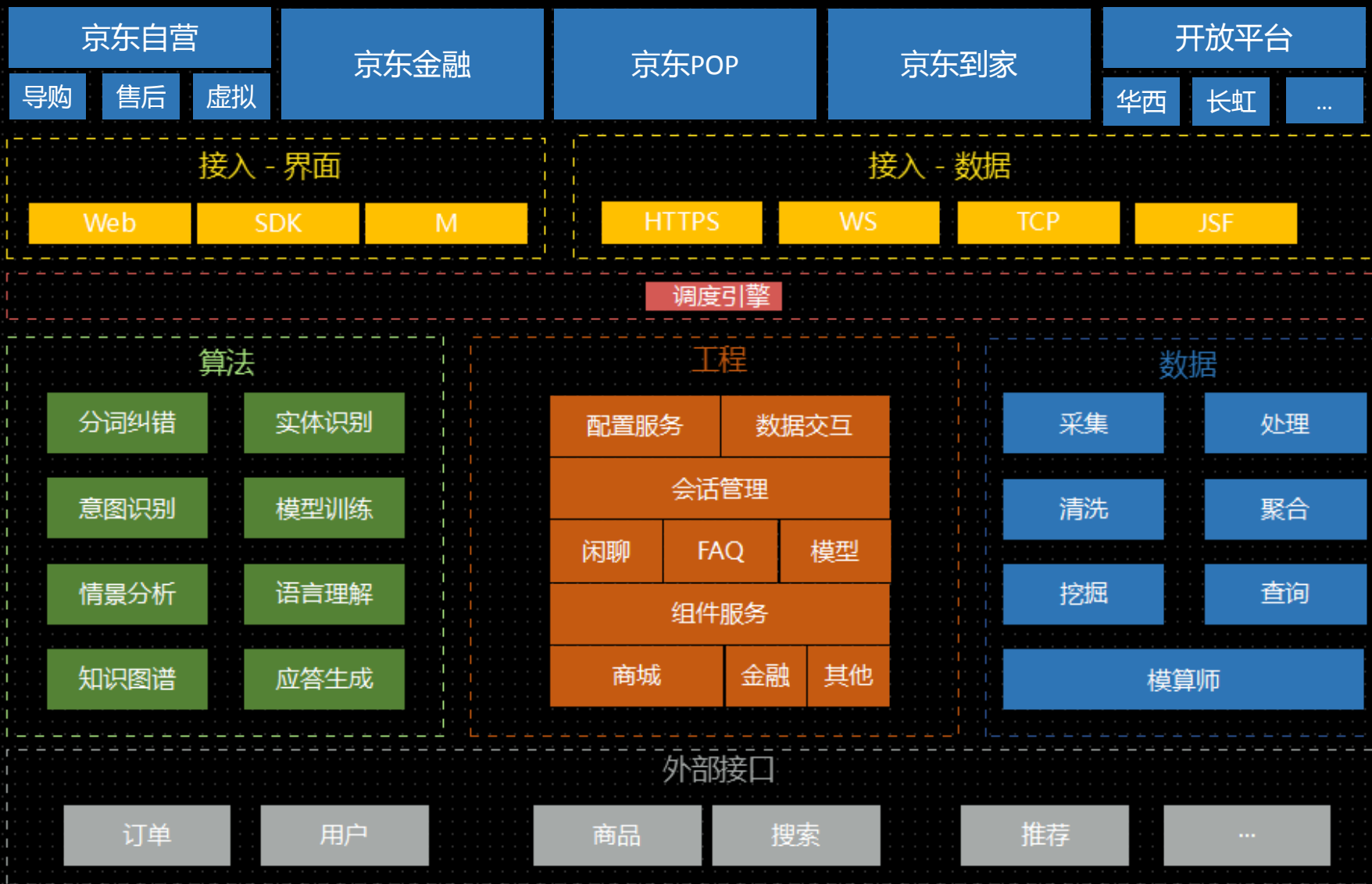
用户情绪感知

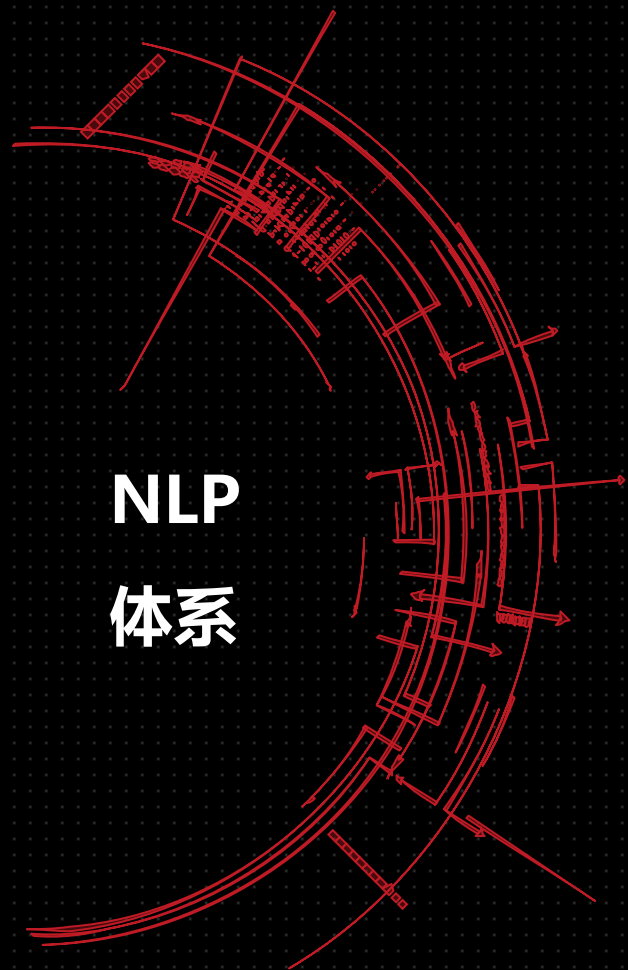
感知用户情绪并安抚

个性化应答

不同客户，不同体验

JIMI机器人核心架构





纠错、分词



实体识别

用户意图识别



句法
分析

深度
神经网络

情绪
识别

应答
引擎

FAQ

用户未来意图预测

- 项目介绍：根据用户的实时浏览、点击、当前问题等信息，智能预测用户未来的意图，并应用在产品中



用户未来意图预测——用户进入咨询前预测

- 场景描述：在用户进入JIMI还未提问时，提前预测用户可能问的问题
- 算法选型：1、**聚类**找到标准问题；2、根据用户点击获取**分类标签**

原始问题收集

按热门SKU维度：

- JIMI聊天日志
- IM聊天日志
- 单页购买咨询
- **LR模型**数据清洗



标准问题构造

- 分词
- word2Vec
- 去停用词，同义词替换
- 取中心词生成句子向量
- 聚类生成标准问题



特征收集与建模

- 接口数据：订单、服务单等
- 实时数据：浏览、点击
- 画像数据：性别、年龄
- 建立分类模型

基于模型数据清洗，快速提升清洗效率

- ✓ 人工标注+规则过滤 -> 基于LR模型数据清洗

技术原理

- ✓ 人工标注获得正负样本
- ✓ LR训练模型

效果

- ✓ 清洗10000样本时间**2人天->5分钟**

增加实时特征，提升模型准确度

- ✓ 获取用户全网浏览、点击实时数据特征

技术原理

- ✓ 特征one-hot编码，PCA降维

效果

- ✓ 增加实时特征**300->20000+**
- ✓ 提升模型准确率**73%->78%**

优化聚类算法，减少聚类时间

- ✓ 传统K-means -> 优化Canopy + K-means

技术原理

- ✓ 设距离阈值 $T1$ ，当任意点 $P \in$ 集合 S ，且 $P < T1$ ， P 从集合 S 移除加入当前Canopy，迭代至 S 为

效果

- ✓ 每1000条语料聚类时间**减少近27s**

分布式任务，降低运行时间

- ✓ 单实例计算 -> 分布式任务

技术原理

- ✓ MapReduce任务处理
- ✓ 基于Spring Batch框架

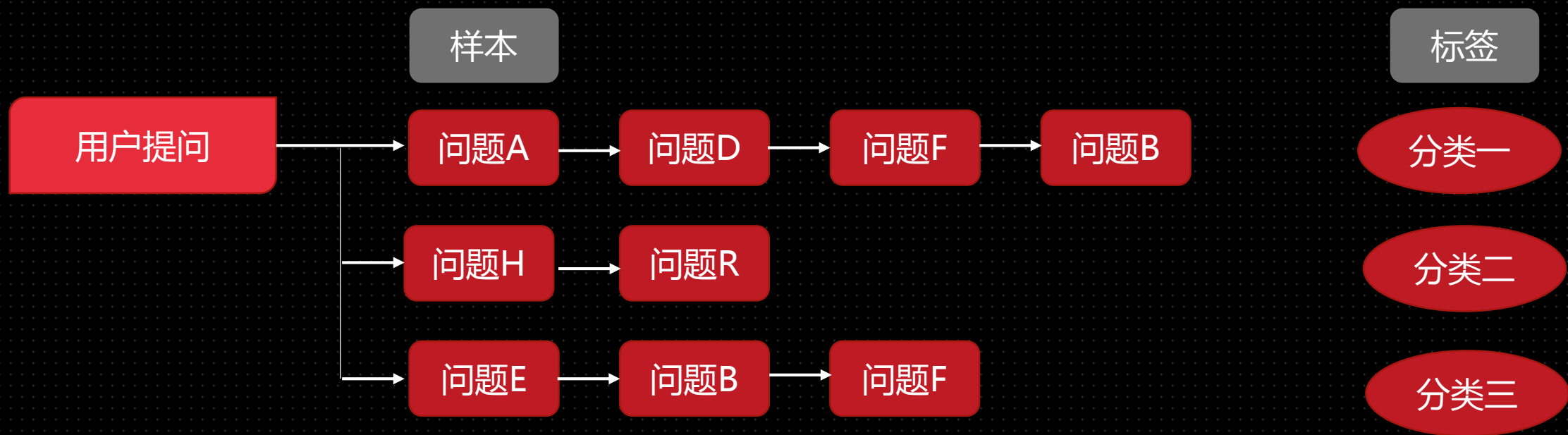
效果

- ✓ 降低全流程运行时间**8h->1.5h**

用户未来意图预测——用户咨询过程中实时预测

- 场景描述：在用户进入咨询过程中，根据用户当前问题，实时预测用户未来意图，并将预测的意图展示给用户点选，提升用户体验
- 算法选型：基于LSTM的深度学习模型

建模思路：



用户未来意图预测——用户咨询过程中实时预测

模型调优经验

- ✓ L2正则化
- ✓ 样本 > 1000000条
- ✓ 加入Dropout机制
- ✓ 截断式反向传播
- ✓ 掩模和填零
- ✓ 学习率自动衰减：AdaGrad
- ✓ 权重初始化：XAVIER
- ✓ 优化方法：SGD/AdaGrad/momentum
- ✓ 激活函数：softmax

模型效果评估

- ✓ 建立BaseLine：针对当前分类X，基于历史数据统计，给出最高频的下一分类Y

最终效果数据

模型	点击率	准确率
Baseline	52.4%	62.4%
LSTM	68.7%	75.5%
LSTM + Attention	71.2%	78.2%

JIMI+咚咚 = 无人客服

» 新一代智能客服 «



第四代智能客服引擎

深度神经网络，更智能

咚咚+JIMI深度融合

人机无缝衔接

全流程智能再造

C端/B端全流程升级

人机交互式自学习

高质量大数据沉淀与挖掘



场景预判



JIMI接入

无缝

转换



专家坐席



人工接入

人机无缝深度融合

演示Demo

消除沟通断层感



提升用户体验和咨询效率



2017年每月节省人工成本

上千万元

日最高接待量

数百万次

响应时间

< 100ms

覆盖商品

十亿+

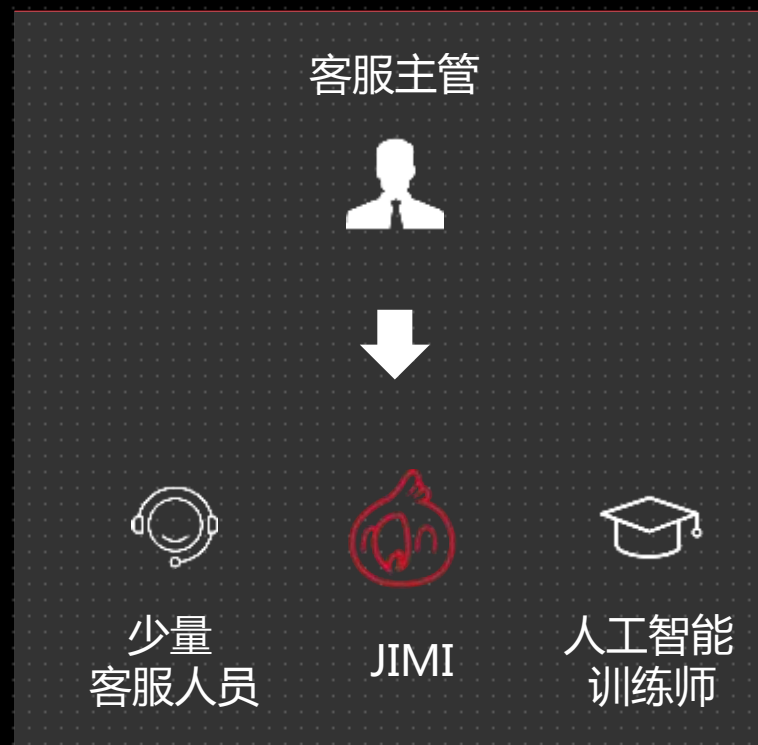
应答准确率

> 90%

用户满意度

> 80%

以京东人工智能之力，改变客服行业





THANK YOU

京东JIMI



本PPT来自讲师在携程技术沙龙上的分享，更多
干货可关注“携程技术中心”微信公号