



NLP项目

课程安排

授课内容

- 【目标】：了解NLP的常用应用场景及系统
- 【掌握技能点】：PyTorch基础框架、Jieba&HanLP&Gensim等各种工具的使用、Seq2Seq&Transformer&Bert&Attention等基础网络结构、分类&信息提取&信息生成等相关应用、大语言模型的基本应用；
- 【预计授课时长】：19~21天(仅考虑周末)
- 【资料网盘链接】：
 - 【连接】：<https://pan.quark.cn/s/82bf0318de0c> 提取码：dmqr
 - 【连接】：<https://www.123865.com/s/kBfwvd-bSWRA?pwd=poki#>

NLP概述

■ Natural Language Processing(NLP, 自然语言处理)

■ **目的：让计算机处理或“理解”自然语言，以执行语言翻译和问题回答等任务；最终来讲就是构建机器语言和人类语言之间的沟通桥梁，实现人机交流为最终目的。**

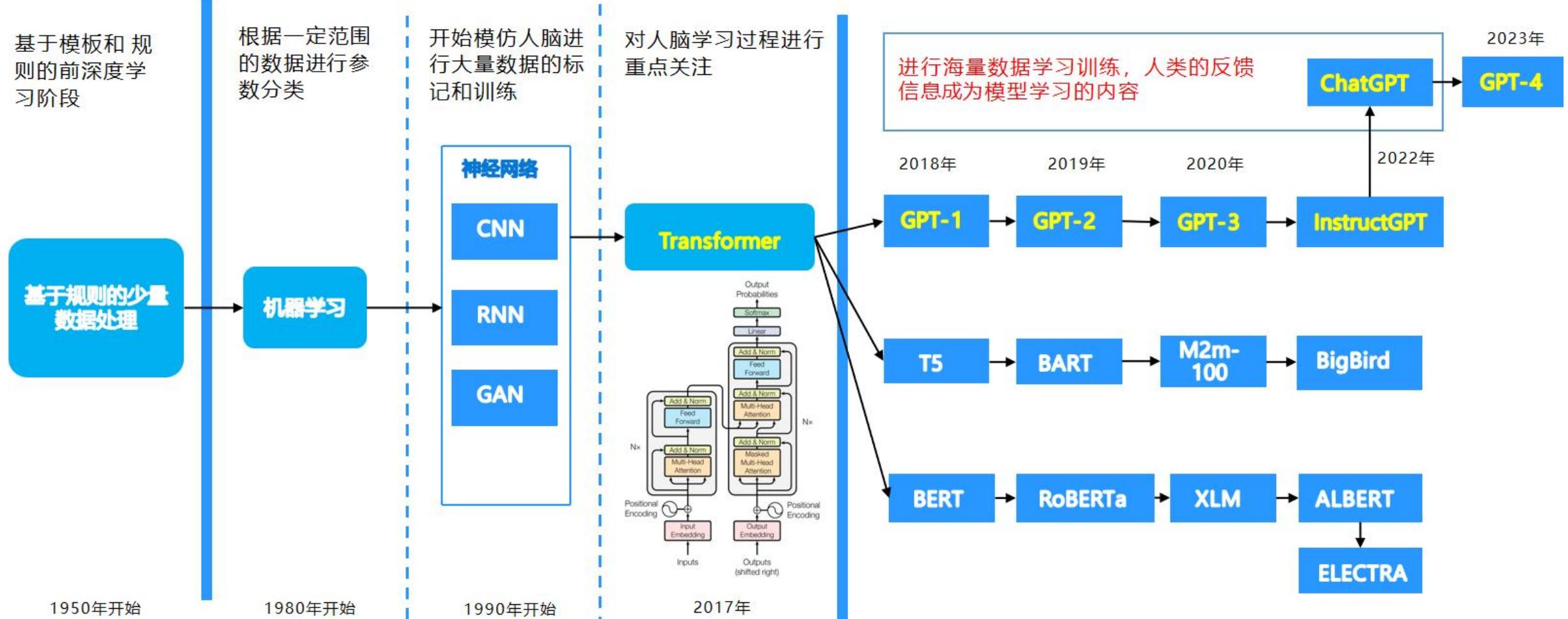
■ 常见应用：

- **抽取：关键词提取、概要抽取、事件提取、命名实体识别(提取价格、日期、姓名、公司、地点等)、关系抽取等**
- **分类：文本分类、情感分析、意图识别、文本相似度、文本向量等**
- **生成：机器翻译、语音文本转换、图像文本转换、文本生成、纠错等**
- **NOTE：自然语言一般来讲模型都是分类模型**

NLP概述

- 💡 从另外一个角度来看，NLP分为NLU和NLG两个子任务：
 - 💡 NLU(Natural Language Understanding)自然语言理解
 - 💡 含义：让计算机理解自然语言文本的意义；
 - 💡 主要任务方向：**分词、词性标注、文本分类、实体识别、信息抽取**等；
 - 💡 NLG(Natural Language Generation)自然语言生成
 - 💡 含义：将非语言格式的数据转换成人类可理解的语言格式；
 - 💡 主要任务方向：**机器翻译、问答系统、聊天机器人、LLM**等；

NLP的发展流程

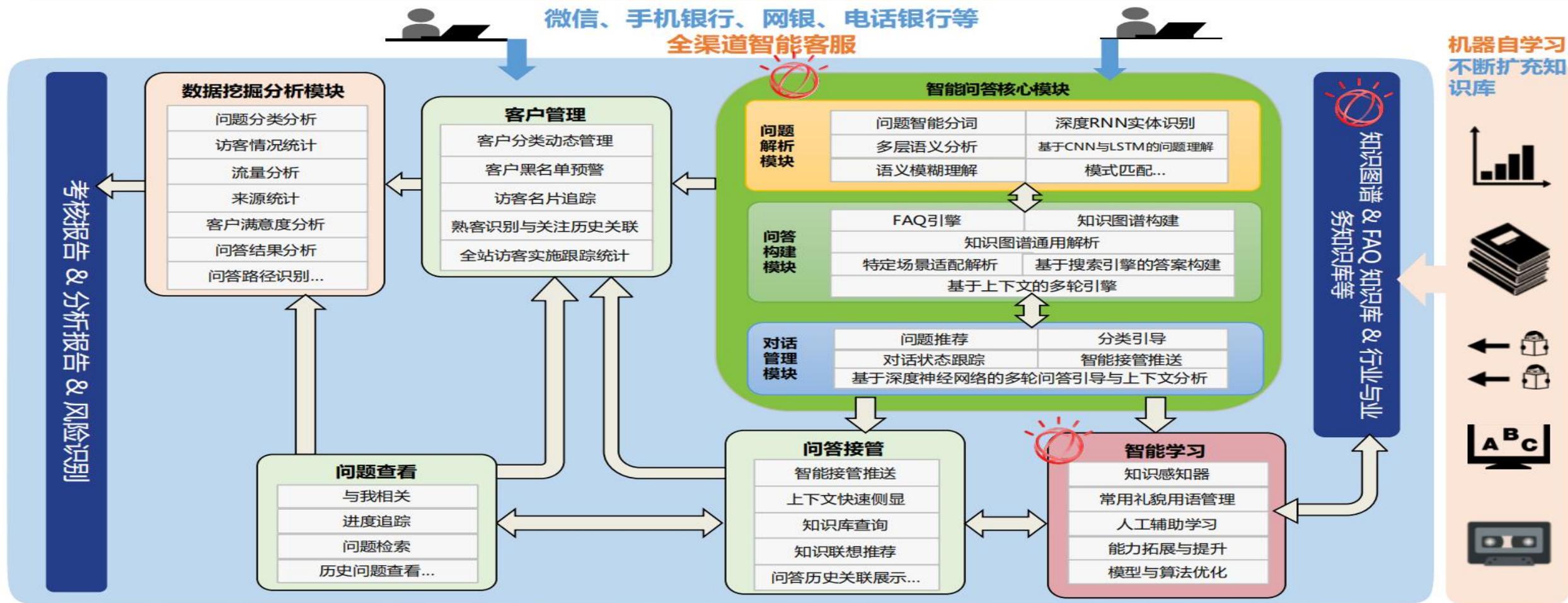


文心一言架构

产品与社区	文心一格	文心百中	旸谷社区						
	AI艺术和创意辅助平台	大模型驱动的产业级搜索系统	大模型创意与探索社区						
工具与平台	EasyDL-大模型 零门槛 AI 开发平台	BML-大模型 全功能 AI 开发平台	大模型 API						
	大模型套件								
文心大模型	数据标注与处理	大模型精调	大模型压缩	高性能部署	场景化工具				
	行业大模型								
国网-百度·文心		浦发-百度·文心	航天-百度·文心	人民网-百度·文心	冰城-百度·文心	电影频道-百度·文心			
深燃-百度·文心		吉利-百度·文心	泰康-百度·文心	TCL-百度·文心	辞海-百度·文心				
NLP 大模型			CV 大模型		跨模态大模型		生物计算大模型		
文心大模型	医疗 ERNIE-Health		金融 ERNIE-Finance		商品图文搜索表征学习 VIMER-UMS		文图生成 ERNIE-ViLG	文档智能 ERNIE-Layout	化合物表征学习 HelixGEM
	对话 PLATO		搜索 ERNIE-Search		OCR图像表征学习 VIMER-StrucTextT				
	跨语言 ERNIE-M		代码 ERNIE-Code		多任务视觉表征学习 VIMER-UFO				蛋白质结构预测 HelixFold
	语言理解与生成				视觉 -语言 ERNIE -ViL	语音 -语言 ERNIE -SAT	地理 -语言 ERNIE -GeoL		单序列蛋白质结构预测 HelixFold-Single
ERNIE 3.0 Tiny (轻量级)	ERNIE 3.0 (百亿级)	鹏城-百度·文心 (千亿级)	ERNIE 3.0 Zeus (任务知识增强)	视觉处理 多任务学习 VIMER-TCIR	自监督视觉 表征学习 VIMER-CAE				

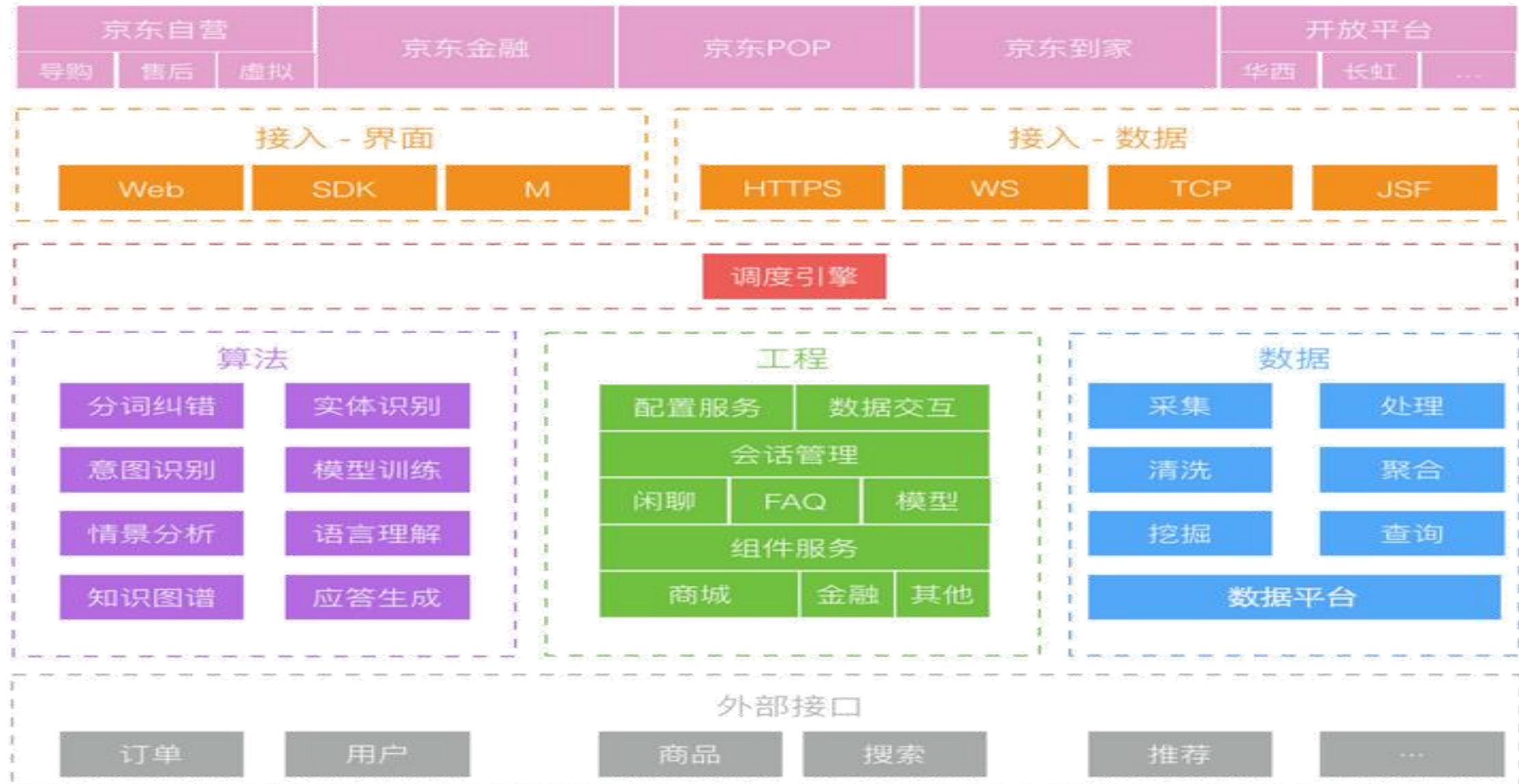
IBM Watson

IBM智能客服—业务咨询功能概览



NLP概述

京东NLP框架



NLP概述

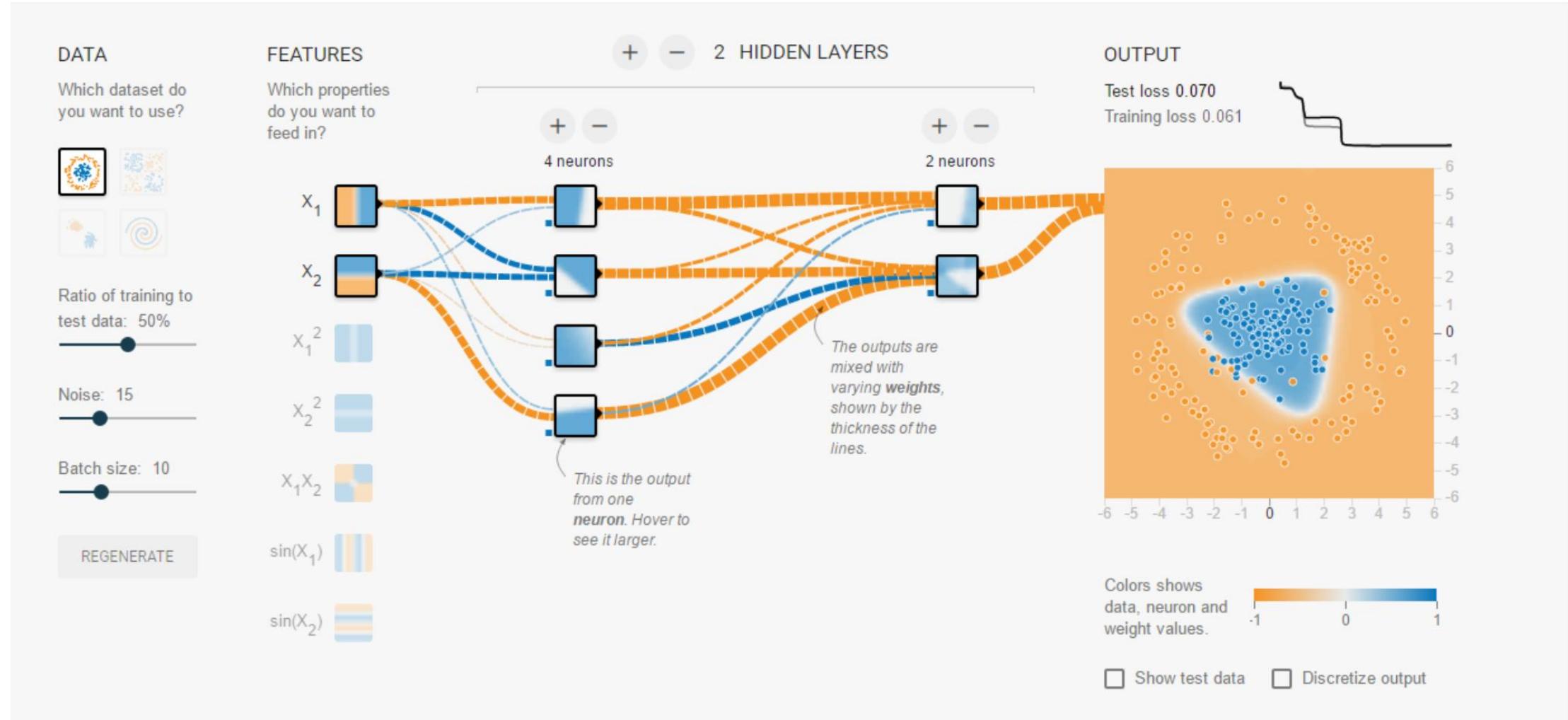
美团点评NLP框架

NLP@美团点评



直观案例

案例理解



THANKS!