
ChatGPT 技术过程和应用发展

大数据与人工智能实战专家—李家贵老师

广东省人工智能与大数据中心技术总监/数据中台部总经理/CDO

课程大纲：

一、ChatGPT 改变世界

1.1 宏观认知

1.1.1 二季度政治局会议关于 chatGPT 会议结论的政策解读

1.1.2 人工智能的 iphone 时刻/登火箭时不要问坐那个座位

1.1.3 AI 的影响可能是反人性的

1.1.4 chatGPT 的影响可能是文艺复兴级的

1.1.5 新的大国竞争和 wintel 联盟

1.1.6 上一轮人工智能进入尾声

案例：5000 万美金的朋友圈

1.2 ChatGPT 的先进性

1.2.1 ChatGPT 具备诸多先进性特征

1.2.2 ChatGPT 提升的核心点

1.2.3 ChatGPT 提升的原因

1.2.4 ChatGPT 提升的领域

1.2.5 ChatGPT 得益于通用（基础）模型所构建 AI 系统的新范式

1.3 ChatGPT 的局限性

1.3.1 机器幻觉

1.3.2 知识库问题

1.3.3 信息安全

1.3.4 跨境传输

第二章：功能分析与应用场景

2.1 测试方向：自动化测试、质量控制等应用。

2.2 数据分析方向（code interpreter）：数据挖掘、预测分析等应用。

2.3 生图分析：图像生成、设计辅助等应用。

2.4 生文分析：各类生文应用。

2.5 生代码分析：代码自动生成、程序优化等应用。

第三章：案例场景探讨

3.1 客服：智能客服系统的构建与优化。

3.2 新媒体营销：内容创作与传播的智能化支持。

3.3 电商运营：运营效率的极致提升。

3.4 财务：财税等应用。

3.5 全媒体营销：知识库+RPA 多平台营销策略与协同应用。

3.6 HR：提单的智能支持。

3.7 本地知识库：本地化知识管理。

第四章：对内赋能（如何用好提示词）

4.1 prompt 的一个本质

4.2 Prompt 的二个心法

4.3 Prompt 的三项沟通要点

4.4 prompt 的四项 chat 法则

4.5 prompt 的五个业务法则

-
- 4.6 Prompt 的六条咒语
 - 4.7 Prompt 的七项注意
 - 4.8 Prompt 的天龙八部
 - 4.9 Prompt 的追问法则
 - 4.10 prompt 的继续法则
 - 4.11 Prompt 的工程化：Auto Prompt，COT prompt，Prompt Perfect
 - 4.12 什么时候该选择 3.5，什么时候该选择 4.0？

第五章：深度前沿-如何预训练和微调

2.1 什么是预训练

为什么要预训练--中文表现不佳

为什么要预训练--领域知识不足

为什么要预训练--有时候微调不解决问题

为什么要预训练--行业实践

预训练的目的

领域差距是什么？

数据问题

预训练过程

分词

词表扩充

数据源采样

中文预训练集

关于数据的结构化

C-Eval

2.2 什么是微调

Instruct tuning

Prompt tuning

Instruct-GPT paper

Alpaca 微调过程

Vicuna-13B 的训练、评估和推理

微调国内实践

数据处理

使用 RLHF 来优化模型

模型优化思路

通过引入判别模型代替人工打分

基于强化学习 RL 进行语言模型微调优化

人类反馈的强化学习 RL 的总流程

大实验：利用竹间智能平台一起共创一个微调政务大模型(0.5天-1天)