# 推荐系统如何从大语言模型中取长补短：从应用视角出发

唐睿明 ---- 华为诺亚方舟实验室

DataFunSummit # 2023

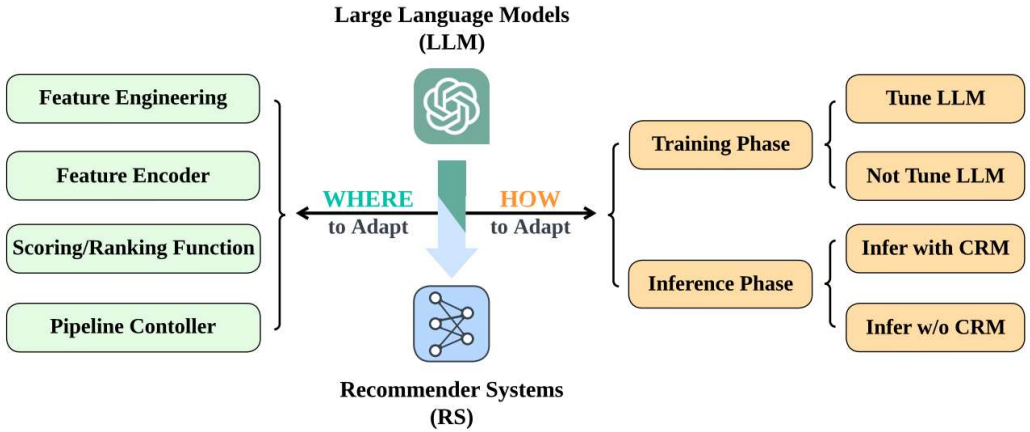
# 目录
# CONTENT

# 01

# 背景和问题

# 背景和问题

- **传统的推荐系统**
  - 模型相对较小，时间空间开销低√
  - 可以充分利用协同信号√
  - 只能利用数据集内的知识×
  - 缺乏语义信息和深度意图推理×
- **大语言模型**
  - 引入外部开放世界知识，语义信号丰富√
  - 具备跨域推荐能力，适合冷启动场景√
  - 协同信号缺失×
  - 计算复杂度高，难以处理海量样本×
- **核心研究问题**
  - 推荐模型如何从大模型中取长补短，从而提升推荐性能，优化用户体验?
- 从应用角度出发，我们进一步将该问题拆解为
  - **何处运用大语言模型（WHERE to adapt）**
  - **如何运用大语言模型（HOW to adapt）**




**LLM+RS：核心研究问题拆解**

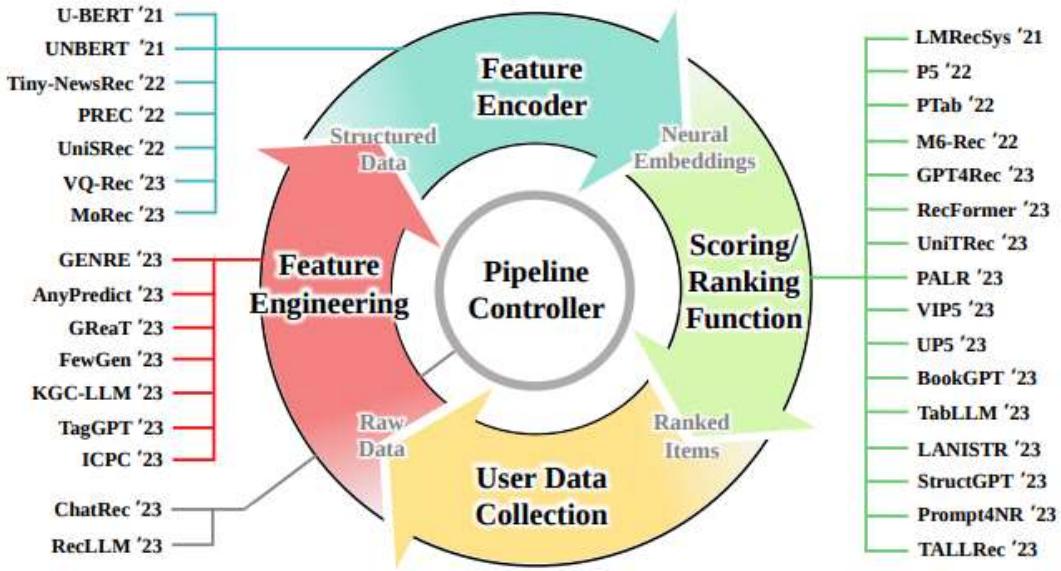
Lin J, Dai X, Xi Y, et al. How Can Recommender Systems Benefit from Large Language Models: A Survey. arXiv preprint, 2023.

# 02
# 何处运用大语言模型

DataFunSummit # 2023

# 何处运用大语言模型 (WHERE to adapt LLM)

- 根据现代基于深度学习的推荐系统的流程，我们抽象出以下五个环节：
  - **数据采集**阶段：线上收集用户行为和记录，得到原始数据（raw data）
  - **特征工程**阶段：对原始数据进行筛选、加工、增强，得到可供下游深度模型使用的结构化数据 (structured data)
  - **特征编码**阶段：对结构化数据进行编码，得到对应的稠密向量表示 (neural embeddings)
  - **打分排序**阶段：对候选物品进行打分排序，得到要呈现给用户的排序列表( recommended items)
  - 推荐**流程控制**：作为中央控制器，把控推荐系统的整体流程。也可以细化到对排序阶段的召回、粗排、精排的控制

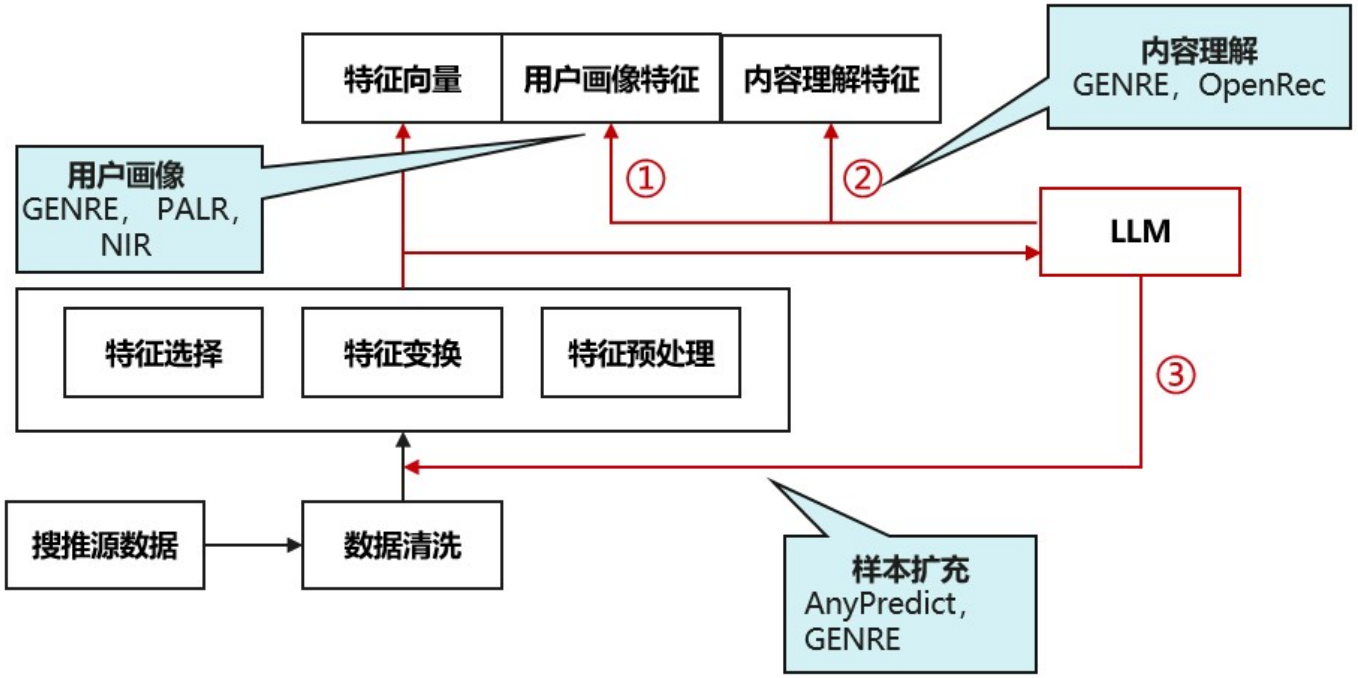



**基于深度学习的推荐系统流程和不同阶段LLM应用的代表性工作**

Lin J, Dai X, Xi Y, et al. How Can Recommender Systems Benefit from Large Language Models: A Survey. arXiv preprint, 2023.

# 特征工程

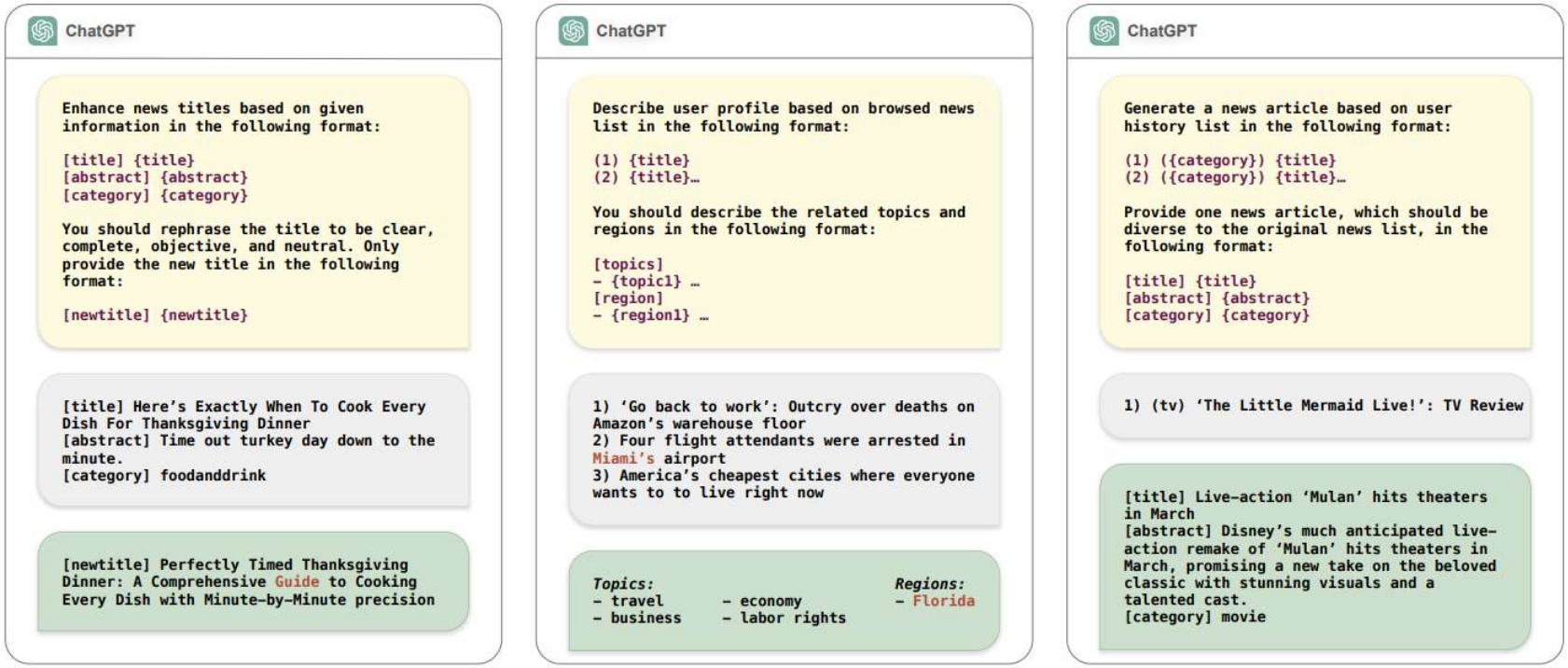
- 利用大语言模型的**外部通用知识**和**逻辑推理**能力，进行特征增强
  - 1. 丰富用户画像 2. 理解推荐内容 3.样本扩充




**大语言模型在特征工程中的应用**

# 特征工程

- GENRE
  - 在**新闻推荐**的场景下，利用LLM进行新闻摘要，用户画像和个性化新闻内容生成

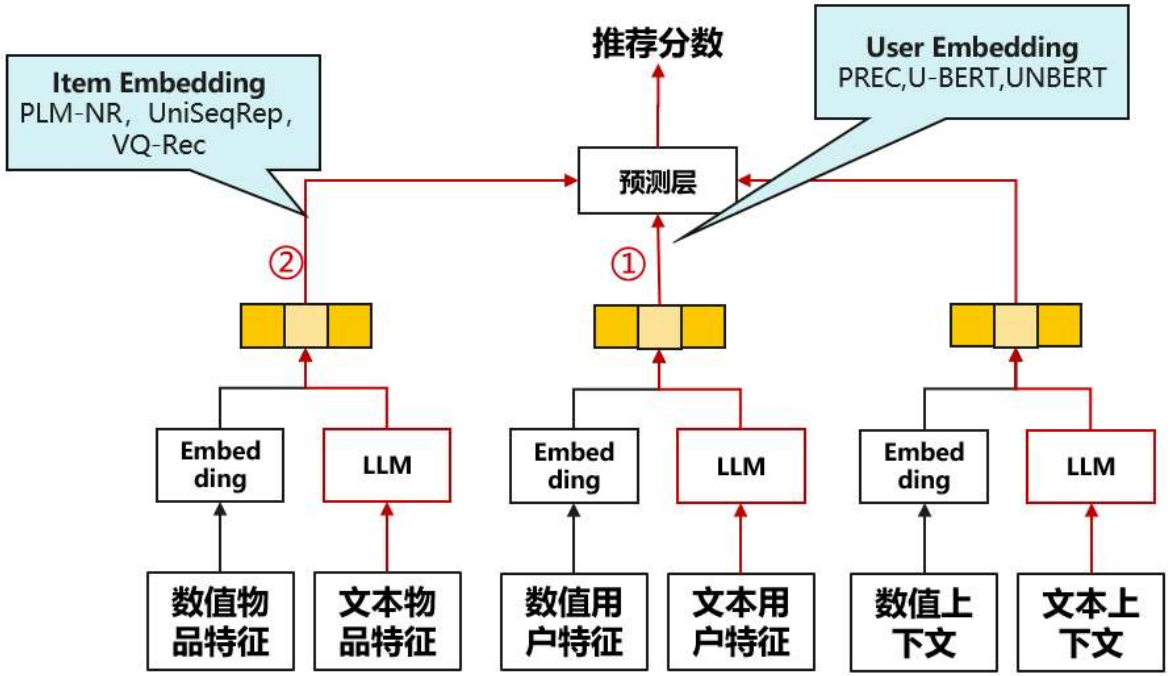



(a) News summarizer　(b) User profiler　(c) Personalized news generator

Liu Q, Chen N, Sakai T, et al. A First Look at LLM-Powered Generative News Recommendation. arXiv preprint, 2023.

# 特征编码

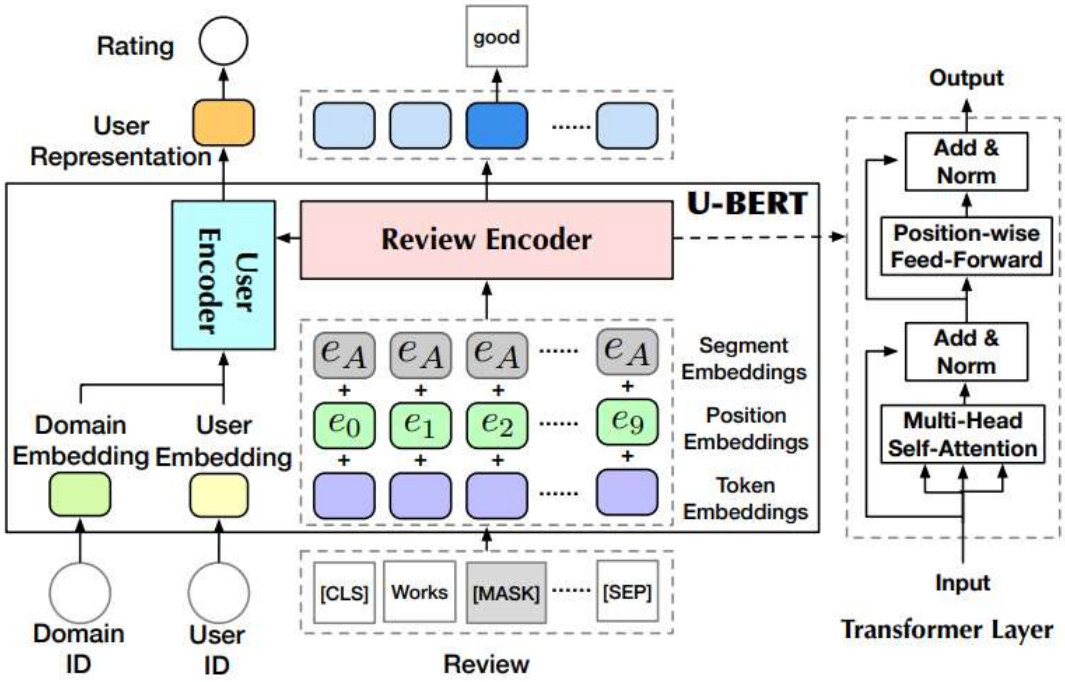
- 利用LLM的通用语义信息丰富推荐特征表示
  - 1.增强文本特征 (用户表征、物品表征) 表示 2.改善基于ID的特征表示的跨场景迁移能力




**大语言模型在特征编码中的应用**

# 特征编码

- U-BERT
  - **用户表征**：用语言模型对用户评论内容编码，增强用户的个性化表征




Qiu Z, Wu X, Gao J, et al. U-BERT: Pre-training user representations for improved recommendation. AAAI, 2021.