ChatGPT原理

- 优缺点和特性
  - 优点
    - 处理任一领域的NLP任务
    - 减少标注量，少量的优质标注就可以辅助他完成任务
  - 缺点
    - 幻觉严重
    - 静态的预训练模型，本质上不具备思维和推理能力
- 强化学习
  - 要素
    - 状态：环境和模型共同构成的整体状态
    - 动作：模型可以做出的动作，比如落子/跑跳等，所有动作可以组成一个有限动作集
    - 策略：状态到动作的映射
    - 反馈：做出决策后，处于下一个状态，并得到反馈
  - 学习路径
    - 状态-策略-行动-反馈 是一个流程
    - 在状态s下，经验和行为只依赖当前状态的性质，这被称之为马尔科夫性
  - 价值函数
    - 状态价值函数，反馈当前状态的期望
    - 行动价值函数，反馈行动对目标的期望
  - 在NLP领域的难点
    - 环境难模拟
    - 价值难定义，人工标注成本又高
  - chagpt的强化学习
    - 基于transformer预训练llm
    - 训练监督模型SFT，微调llm
    - 训练奖励模型，评判sft响应
    - 基于ppo，调整sft
      - on-policy，以当前状态和行为做训练
      - off-policy，以历史状态和行为做训练
      - 损失函数定义为两者差值，期望是两者结果相近
  - 数据集
    - raddit
    - books - 提取的图书文本
    - Wikipedia
    - gpt的能力是超出自然语言的范畴的，还可以理解图形符号，二进制码，代码等
    - 幻觉是个大问题
- 训练方式
  - 纯监督学习
    - 需要大量的样本和高质量标注
    - 只针对特定任务
    - 泛化能力差
    - 基于完全随机的参数值
  - 预训练+微调
    - 基于已训练过的权重进行调整
  - 小样本学习
    - in-context 训练：基于上下文的训练，优化特定场景的表现
    - zero-shot：不给示范对话
    - few-shot：给少量示范对话
    - one-shot：给一个样本
  - prompt学习
    - 直接在对话中告知任务
  - 强化学习RLHF
    - 模型自主尝试，吃一堑长一智
- 编码器解码器
  - gpt保留了解码器，使用transfomer实现
  - transformer的输出与词表中token数量一致
  - 做了正则化，归一化后的transformer输出结果，通过检索规则将token解码为字符
    - 贪婪搜索，直接找概率最大值，容易错过正确答案
    - 束搜索，每一束都挑概率比较高的token，做概率选择
    - 核搜索，从大到小找到一堆概率之和想加小于top_p的token集合，再从中按照不同概率挑选
      - gpt使用温控+核搜索的形式，转换完概率后用top_p抽取，这两个参数都在接口里
    - 温控搜索，用一个softmax公式转换结果概率
- 发展脉络
  - 基于规则：手把手替换和实现规律
  - 基于统计：半引导
  - 基于强化学习：主要靠自学
- 版本特点
  - 1.0 语言编解码，补全上下文
  - 2.0 多任务学习者
  - 3.0 小样本学习，大模型中的大模型
  - 4.0 多模态的大模型
- 模型训练
  - 根据上文猜下文
  - 使用大量语料作为目标，反向调整字的向量
  - gpt权重在训练完毕后是静态的
- token
  - token是输入的最小粒度单位
  - 由有意义的高频词/词根组成
  - 可编码可解码
  - BPE：按字符算，一个中文算一个token
  - Byte-Level-BPE：按字节算，一个中文三个字节 3token
  - 词表可由统计数据得到
  - 作用
    - 克服长尾效应，低频长词可以拆为词根
    - 支持多语言
    - 上可解码为字符，下可转为张量为gpt输入；输出过程反之亦然
- embedding
  - 字符和模型的中间层
  - 抽象token的语意
  - 建立自然语言语意和数字空间的联系
- 注意力机制
  - 移除不重要的部分
  - 专注与自身相关的部分
  - 寻找上下文关系就需要抓到重点
  - 信息压缩
  - 自注意力机制
    - 注意力权重来自于句子本身
    - 好处在于跨越token找到重点
    - 相比rnn，记忆可以更长更重点的影响后续的结果
- transformer
  - 稀疏transformer
    - 本质是为了减少注意力和token的计算量，忽略部分计算
    - 采用 mask 机制屏蔽部分不需要参与注意力计算的token
    - 多头注意力机制：多做几次注意力，让模型关注到不同的信息，常量级别的复杂度
    - normalization正规化：把分布范围缩小，避免过拟合
    - dropout机制：对某些结果随机置零以简化模型，避免过拟合