

Goushi GPT V3 开发思路

我们采用的模型思路是 GPT 风格的模型，也就是类似 ChatGPT 的那种结构。

这种结构叫做“decoder-only transformer”，此处只作了解即可。

这个模型的本质是根据上文预测下一个 token（词或者字，或者其他分割方法，下文会提及）是什么，它通过概率预测，每次随机选择概率比较高的那几个 token 中的一个。

1. 输入处理

模型需要概率预测，那么它肯定不能直接处理我们输入进来的文字。因为数学计算需要数字（废话），所以我们需要先把文字转成数字

这里需要从 token 开始讲起

首先输入需要通过一个叫做 tokenizer 的东西，这个东西会把输入分割成小块，这些小块可能是单个字符、词语，或者其他组合，例如词根词缀的形式。这些小块就是 tokens

tokenizer 会有一个词汇表，它包含了所有的 token，以及它们的对应编号。随后 tokenizer 就会把分割后的 tokens 都转成编号的形式

此时我们成功把输入句子转成一串数字了。

这个部分需要 1-2 人完成

2. Embedding

在模型中，embedding（嵌入）的作用是将离散的符号（如单词、子词或字符）转换为连续的、低维的向量表示，使得模型能够处理和理解这些符号之间的语义关系。

我们用 tokenizer 把句子转成一串数字之后，训练一个多层感知器（MLP），把这些数字作

为输入给这个 MLP，让它来把我们的这对数字转成连续低维向量表示。

这个部分需要 1 人完成

3. Decoder

这是难度最高的部分。

在 GPT 风格的模型（即基于 Transformer 架构的纯 Decoder 模型）中，Decoder 的核心作用是根据已有的输入序列，自回归地生成下一个词（token），从而完成文本生成任务。

具体来说，GPT 中的 Decoder 具有以下几个关键功能：

自注意力机制（Self-Attention）：

Decoder 通过自注意力机制捕捉输入序列中各个词之间的依赖关系。由于 GPT 采用因果掩码（causal masking），每个位置只能关注到它之前（包括自身）的词，确保生成过程符合时间顺序，防止“偷看未来”。

逐词生成（Autoregressive Generation）：

GPT 以自回归方式工作：给定一个起始提示（prompt），模型依次预测下一个最可能的词，并将新生成的词加入输入序列，重复此过程，直到生成完整输出。

无 Encoder 结构：

GPT 不包含 Encoder 部分，仅使用 Decoder 堆叠多层，专注于从左到右的单向语言建模任务（即预测下一个词的概率分布）。

语言建模目标：

Decoder 在训练时的目标是最大化整个序列的联合概率，即通过极大似然估计学习条件概率：

$$P(w_1, w_2, \dots, w_n) = \prod_{t=1}^n P(w_t | w_1, \dots, w_{t-1})$$

这个部分需要 1 人完成

4. 模型的训练代码

我们的模型采用监督学习。因为它的本质任务就是不停预测下一个 token 是什么，训练时使用的 label 是输入序列中每个 token 的下一个 token。

比如：

对于数据集中输入的句子：

The cat sat on the mat.

label 的情况如下：

输入 (context)	Label (next token)
"The"	"cat"
"The cat"	"sat"
"The cat sat"	"on"
"The cat sat on"	"the"
"The cat sat on the"	"mat"
...	"."

我们需要写出代码来完成这个。

这个部分需要 1 人完成

5. 模型的调用

调用训练之后的模型！

我们采用一个名为 Gradio 的 Python 库，它可以让我们快速地搭建一个与模型交互的界面。

需要在这里调用我们的模型，输入模型的一些参数

这个部分需要 1 人完成

6. 额外工作：数据收集

既然要训练模型，我们需要收集我们的训练数据，就像课上讲的一样，模型需要在已有数据上“刷题”，学习数据的模式，尝试接近我们最希望它输出的结果，自动地调整各项参数，这个叫做训练，训练之后才可以使用。

我们这个模型是中文生成模型，需要收集一些中文数据集，例如小说、对话等。而且，我们的模型以抽象猎奇作为特点，所以这些数据最好是猎奇一些的文案

示例：

今天我终于狠心买了一瓶矿泉水尝尝，能喝上矿泉水一直是我的梦想。每当我在大街上看到喝矿泉水的人，我都会偷偷瞄一眼，真羡慕那些有钱人。昨天我发了压岁钱，在经过激烈的思想斗争之后，我终于下定决心明天去买瓶矿泉水喝，我没有告诉我的父母，他们知道后一定会阻拦我的，我激动的一宿没睡。第二天，我花了2个小时半来到市中心的矿泉水专卖店，我完全被那种气势震住了，豪华的装修，我怀着忐忑不安的心情走进了专卖店，面积达8平米的营业大厅布置和外面一样豪华，友善的服务员热情的接见了我，可能是太过紧张了我开口要求试吃，刚说完我就脸红，可是服务员并没有感到诧异，对了，可能来买矿泉水的人都得要求先试喝吧，毕竟2元/瓶的价格不是说买就买的，我激动地接过矿泉水，小心翼翼地拧开了瓶盖，这时的我已经忘了要注意绅士风度，贪婪的喝了一口，爽滑的口感不是我以前喝的自来水所能比拟的。美丽漂亮的女服务员热情的向我讲解矿泉水的配料和制作流程，在谈话中我得知她刚从韩国首尔大学毕业，主修酒店管理双硕士学位，回国后她拒绝了去星巴克当店长的机会，来到矿泉水专卖店工作，我也为她的明智选择而感到欣慰，毕竟进入矿泉水专卖店是多少代人的梦想与追求。临走时我掏出带着体温的4元钱买了2瓶，刚要走出店门时服务员朝我喊：“你的矿泉水！”我回眸一笑百媚生回应道：“是你的矿泉水。”服务员当时就脸红了，我随即以迅雷不及掩耳盗铃儿响叮当之势走出门口。我想，服务员肯定会把所有的矿泉水从柜子上都拿下来一瓶一瓶的数“他明天会来，他明天不会来……”走在大街上我喝着矿泉水，大家都向我投来羡慕的目光，我也感觉对人生更加有信心了。回到家，我兴奋得不能自己，还是忍不住和我的矿泉水合照了，因为矿泉水实在太过耀眼，我怕有些不法分子会盯上我，所以我用美丽的签名挡住了我俊朗的面庞。嘘！记得不要告诉别人我喝过矿泉水，这年头像我这样低调的人不多了。

B 站上有很多这种抽象文案的视频，你可以用浏览器把它们下载下来，然后直接发给我处理就好了。记得标明出处。

从其他地方复制粘贴也可以

注意版权问题，有的来源我们需要在我们模型介绍里面注明。所以一定要比标记出处。

这个部分需要 1-2 人完成

附录：模型运作原理示意图

