

技术报告

智能问答是自然语言处理的下游任务之一，旨在根据用户的提问生成流畅准确的回复。根据赛题的描述可知，本题目希望将问答系统应用于金融行业，并利用检索技术增强答案的生成。

针对具体的题目要求，我们采用迁移学习的思想，利用特定的数据对预训练模型进行微调，将其迁移至金融领域。同时，我们将检索出来的数据融入原始的数据集，探索检索技术对模型性能的影响。具体的方法和模型概述如下所示。

方法概述

预训练模型在近年来得到极大的关注，被广泛应用于机器翻译，对话系统，智能问答等 NLP 任务。在本赛题中，我们以预训练的中文 `bart` 为基座模型，在特定的数据集上进行微调，将其应用于下游任务。实验结果表明，微调后的 `bart` 可以在一定程度上回答用户的问题，且模型的鲁棒性较强，学习率在 10^{-5} 至 10^{-5} 之间模型的性能差别较小，但在 `rough`、`bleu` 指标上偏低。因此，我们尝试利用检索技术增强模型理解和生成的能力。

假设原始的训练数据为 D ，其包含若干个问题-答案对 $\{query_i: answer_i\}$ 。基于检索增加的思想，对于每一个问题 $query_i$ ，我们首先从候选集合¹中挑选出与问题 $query_i$ 相似程度²最高的问题 $query'_i$ ，之后将其对应的答案 $answer'_i$ 拼接到原问题的答案后面，构造出检索增强的训练数据 D' 。

模型概述

我们采用在中文数据集上预训练过的 `bart` 作为基座模型³。`BART` 吸收了 `BERT` 的 `bidirectional encoder` 和 `GPT` 的 `left-to-right decoder` 各自的特点，建立在标准的 `seq2seq Transformer model` 的基础之上，这使得它比 `BERT` 更适合文本生成的场景；相比 `GPT`，也多了双向上下文语境信息。在生成任务上获得进步的同时，它也可以在一些文本理解类任务上取得较好的结果。

¹ 候选集合即为与原问题最相关的问题的集合，对应赛题中的 `retrieval.search_qa.question` 字段

² 对应赛题中数据的 `retrieval.search_qa.score` 字段，表示与原问题的相似程度，分数越大，相似程度越高

³ [fnlp/bart-base-chinese · Hugging Face](https://huggingface.co/fnlp/bart-base-chinese)

具体来说，我们使用 `huggingface` 作为实现的主要工具，通过相应接口调用其中的 `fnlp-bart-base` 模型。

实验结果

我们以 `pytorch` 为主要的算法框架，并使用 `huggingface` 库实现预训练模型的调用。在实际的实验过程中，受制于服务器的硬件资源，我们采用梯度累计的方式达到增大 `batch size` 的效果。具体实验硬件环境如下表所示：

设备号	设备类型	设备容量
0	TITAN RTX	24220MB
1	TITAN RTX	24220MB
2	TITAN RTX	24220MB
3	TITAN RTX	24220MB

我们使用上表中四张 `gpu` 并行微调中文 `bart`，将每一个设备的批量大小(`batch size`) 设置为 16，梯度累计步数 (`gradient accumulate step`) 设置为 32，学习率为 5×10^{-5} ，最大训练轮数 (`epoch`) 为 5。

经测试，采用上述的实验，模型的最优性能如下所示⁴：

Score	Blue	Blue1	Blue2	Blue3	Blue4
1.14	1.14	6.55	1.27	0.53	0.38

c-rough1	c-rough2	c-roughLsum	c-roughL	meteor
24.00	11.49	19.06	19.23	0.22

⁴ 由天池平台测试