

关于 OpenGPT 的调研

在我们讲述 transformer 为基础的预训练语言模型时，有一个模型我们并没有讲述，这就是 OpenGPT。之所以省略到 OpenGPT 主要是由于该模型在相当多的情况下，被认为是 BERT（以及其他升级版预训练语言模型）的“陪跑版”。

近来, <https://arxiv.org/abs/2005.14165> 一文提出了 OpenGPT v3。

在这种情况下，我们不妨研究一下 OpenGPT 的一些性质。研读上述论文及相关论文，并尝试回答以下问题（一些问题是开放式问题）。

1. OpenGPT 和 BERT 的区别在哪里？
2. OpenGPT 是否在一些任务上（除文本分类外）有一些更好的表现？可能原因是什么？
3. 在该论文中，OpenGPT v3 的主要优势在哪里？
4. 针对该论文的结论以及实用性，你有什么想法？