

# AlphaFold2极简总结

 <https://zhuanlan.zhihu.com/p/390718242>

None

Sat Jul, 24 22:32

简单看了一下AlphaFold2的Nature文章研究了一下设计思路，初步总结如下：

1. 训练神经网络来对regression target进行逐步迭代精化（Iterative refinement）确实很有用。非常多的文章用过这个思路，比如说计算机视觉里的姿态估计（pose estimation）。
2. 广泛运用了Attention架构。一个二维的表可以横着做再竖着做attention，一个图可以在各种局部结构上做attention，从而不断精化embedding的结构。最后那个structure module生成3维结果还是用的attention（invariant point attention），其背后的设计逻辑是对于一个非常复杂的图结构，只找到最相关的结点来进行精化，这样可以降低样本复杂度。
3. 用带标签数据（氨基酸序列与三维坐标的对应）先训练一遍网络，然后用训练完的网络在无标签数据（仅有氨基酸序列）上预测一遍生成新的数据集，只保留预测得好的部分，然后把这两者混合拿来再进行训练，效果更好。这个是之前Google Brain用过的noise-student的训练方案，现在用到这里来了。
4. 类似BERT的masking操作，对各种输入信息加噪音要求输出稳定。这样提高鲁棒性和泛化能力。这个可以说是常用操作了。

总的来说是站在了过去大量神经网络训练成功实践的基础之上，再加上大量的算力和长达几年的专注得到的结果。

最后夹带一下私货：第一条在我们在NeuIPS19的[NeuRewriter](#)和[Coda](#)上也用过，对于离散优化问题相当有效。而在我们最近的一篇做program synthesis的[文章](#)里也用到了与第三条相似的方案：从随机生成含有大量冗余的程序集开始，先训练出一个模型，然后用这个模型做预测会得到一个更紧凑的，去除冗余的程序集，再用这个程序集再次训练模型，效果会变好。

以上。

发布于 07-18

## 文章被以下专栏收录



公号[机器学习算法与自然语言处理] 微信号yizhennotes