

知识点总结

作者姓名

2024 年 8 月 14 日

摘要

模型加速; 大模型; 语音; 图像; 多目标跟踪等。

1 大模型相关技术

1.1 Superalignment

1.1.1 Deep Reinforcement Learning from Human Preferences, 2023

这篇题为《从人类偏好中学习的深度强化学习》的论文探讨了通过引入人类反馈来改进强化学习 (RL)，以在没有预定义奖励函数的情况下优化学习过程。以下是论文的详细解读：

a. 引言

- **问题陈述：**传统的强化学习依赖明确的奖励函数来实现目标。然而，在许多实际场景中，这样的奖励函数往往非常复杂、不明确，或难以制定。这限制了强化学习在这些领域中的应用。
- **解决方案：**论文提出通过使用人类的偏好来定义强化学习代理的目标。通过收集人类对不同结果（轨迹片段）的比较反馈，RL系统可以学习优化符合人类偏好的行为。

b. 方法论

- **从人类反馈中学习：**该方法并不是直接使用奖励信号，而是基于人类偏好学习奖励函数。人类被要求比较代理行为的短视频片段，然后这些比较结果用于训练奖励预测模型。
- **策略优化：**一旦从人类偏好中估计出奖励函数，就可以使用传统的RL算法（如优势演员-评论家

算法(A2C)和信任区域策略优化(TRPO))来优化策略。

- **偏好获取：**人类监督者会看到代理行为的短片段，并被要求指出他们更喜欢哪个片段。这些偏好随后用于更新奖励函数。
- **选择查询：**该方法通过关注那些对正确偏好存在高度不确定性的轨迹片段来选择人类反馈，从而减少所需的比较数量。

c. 实验结果

- **领域：**实验在两个领域中进行：**Atari**游戏和使用MuJoCo的模拟机器人任务。任务的复杂性各不相同，展示了该方法在不同环境中的有效性。
- **性能：**该方法能够以最少的人类反馈解决大部分RL任务。在某些情况下，代理的表现与使用显式奖励函数训练的代理相当，甚至超过后者。
- **新行为：**该方法成功地训练代理执行新颖的复杂行为，如模拟机器人中的后空翻，这些行为如果使用传统的奖励函数来指定是非常困难的。

d. 讨论与结论

- **可扩展性：**研究表明，人类反馈可以扩展到复杂的RL任务，使其在实际应用中变得可行。
- **效率：**这种方法显著减少了所需的人类反馈数量，达到几个数量级的减少，使其更具经济性。
- **未来工作：**作者建议进一步提高从人类偏好中学习的效率，并扩大该方法可以解决的任务范围。

e. 附录

- **实验细节：**附录提供了实验设置的技术细节，包括使用的具体算法、超参数以及提供人类反馈的承包商所接受的指令。

1.1.2 Reward Learning from Human Preferences and Demonstrations in Atari, 2018	1.4.2 RAG: Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks, 2020
1.1.3 Align Before Fuse: Vision and Language Representation Learning with Momentum Distillation, 2021	todo
todo	1.5 推理加速
1.1.4 Fine-Grained Semantically Aligned Vision-Language Pre-Training, 2022	1.5.1 Knowledge Distillation: A Survey, 2021
1.1.5 Scaling Up Visual and Vision-Language Representation Learning With Noisy Text Supervision, 2021	todo
1.1.6 PyramidCLIP: Hierarchical Feature Alignment for Vision-Language Model Pretraining, 2022	1.5.2 LayerDrop: Structured Dropout with Conditional Computation, 2019
1.1.7 FLAVA: A Foundational Language And Vision Alignment Model, 2022	todo
1.2 RLHF	1.6 模型压缩与量化
1.2.1 Fine-Tuning Language Models from Human Preference, 2017	1.6.1 Distilling the Knowledge in a Neural Network, 2015
todo	todo
1.2.2 Learning to Summarize with Human Feedback, 2020	1.6.2 Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding, 2016
todo	
1.3 幻觉抑制	todo
1.3.1 Faithful and Controllable Summarization, 2020	1.7 多模态
todo	
1.3.2 Reducing Transformer Depth on Demand with Structured Dropout, 2020	1.7.1 ViLBERT: Pretraining Task-Agnostic Visiolinguistic Representations for Vision-And-Language Tasks
todo	todo
1.4 RAG	
1.4.1 REALM: Retrieval-Augmented Language Model Pre-Training, 2020	1.7.2 Multimodal Transformers: Vision-And-Language Pretraining with Limited Data, 2020
todo	todo

1.8 Openai Papers

1.8.1 GPT系列-Improving Language Understanding by Generative Pre-Training, Gpt1, 2018

1.8.2 GPT系列-Language Models Are Unsupervised Multitask Learners, Gpt2, 2019

1.8.3 GPT系列-Language Models Are Few-Shot Learners, Gpt3, 2020

1.8.4 DALL-E系列Zero-Shot Text-To-Image Generation, DALL-E, 2021

1.8.5 DALL-E系列Hierarchical Text-Conditional Image Generation with CLIP Latents, DALL-E2, 2022

1.8.6 CLIP-Learning Transferable Visual Models From Natural Language Supervision, 2021

1.8.7 Reinforcement Learning系列-Playing Atari with Deep Reinforcement Learning, 2013

1.8.8 Reinforcement Learning系列-Learning Dexterous In-Hand Manipulation, 2018

1.8.9 Reinforcement Learning系列-OpenAI Five, 2019

1.8.10 安全性与对齐-Aligning Superhuman AI with Human Intentions: A Bayesian Approach” (Irving Et Al, 2019

1.8.11 安全性与对齐-AI Safety Via Debate, 2019

1.8.12 AI Policy与影响力-AI and Compute” (Amodai Hernandez, 2018)

2 模型加速

2.1 Rethinking the Value of Network Pruning, ICLR, 2019

2.1.1 创新点

- **创新点1:** 论文质疑传统的剪枝方法是否真有必要，尤其是在对网络进行结构化剪枝时，直接训练一个小模型的效果可能并不比传统方法差。
- **创新点2:** 大模型中的重要权重在剪枝后的小模型中可能并不关键，剪枝后的架构本身更为重要。

- **创新点3:** 结构化剪枝不仅用于压缩模型，也可视为一种发现高效网络架构的方式。关于结构化剪枝的介绍见表1

2.2 Neural Network Pruning with Residual-Connections and Limited-Data, CVPR, 2020

2.2.1 创新点

- **剪枝残差连接:** 传统的剪枝方法往往仅剪除残差块内部的通道，而保留输出通道不变。论文提出同时剪枝残差块内部和外部的通道，形成类似“钱包”形状的网络结构。这种结构比传统的“沙漏”形结构更有效，在相同的计算量下能够实现更高的精度和更快的推理速度。
- **有限数据集的剪枝:** 剪枝后的模型通常需要在大量数据上进行微调，以恢复其性能。然而，当可用数据集较小时，微调的效果可能不理想，导致模型性能下降。为解决这一问题，论文提出通过数据扩展和知识蒸馏（knowledge distillation）技术，结合标签精炼方法，来提高小数据集上剪枝模型的精度。

2.2.2 有限数据集剪枝的策略

(1) 数据扩展

- **数据扩展:** 通过对现有数据进行各种变换（如旋转、平移、缩放、翻转、颜色调整等）来生成更多的训练样本。这种方法可以在一定程度上增加数据的多样性，缓解数据不足的问题。
- **增强数据集的代表性:** 通过生成不同变体的数据样本，数据扩展可以帮助模型更好地泛化，从而在有限数据的情况下提高模型性能。

(2) 知识蒸馏

- **教师-学生模型:** 知识蒸馏是一种训练技术，其中一个较大或更复杂的“教师模型”被用来指导一个较小或剪枝后的“学生模型”。教师模型可以提供“软标签”（即概率分布，而不是硬标签），这有助于学生模型在有限数据集上的训练。

特性	结构化剪枝 (Structured Pruning)	非结构化剪枝 (Unstructured Pruning)
剪枝对象	层、卷积核、通道、子网络	个别权重或连接
结构保留	保持网络整体结构	破坏网络结构
硬件加速	易于硬件加速	硬件加速困难
适用场景	实际部署与推理场景	主要用于研究或需要极高压缩率的场景
稀疏性	保持一定的结构性	非结构化、稀疏性更高
剪枝后的复杂度控制	通过删除较大的结构块来减少复杂度	精细控制复杂度，通过个别权重调整

表 1: 结构化剪枝与非结构化剪枝的对比

- **蒸馏损失:** 学生模型不仅学习数据集中的硬标签 (即标准分类标签), 还学习从教师模型输出的软标签。这种方法通过补充信息, 帮助学生模型在有限数据集上更好地学习。

知识蒸馏的基本原理

(a) 训练**教师模型**; (b) 教师模型会对训练数据生成**软标签**, 这些概率分布不仅包含了正确类别的高概率, 还包括其他类别的低概率。这些软标签可以捕捉到样本之间更细微的关系; (c) 将教师模型生成的软标签与标准的硬标签结合起来, 用来训练一个较小或较简单的模型 (学生模型); (d) 分类损失 (学生模型的预测与真实标签之间的损失 (通常使用交叉熵损失)) 和**蒸馏损失** (学生模型输出的概率分布与教师模型的概率分布之间的损失, 通常使用 KL 散度。); (e) 在生成教师模型的软标签时, 通常会使用一个**温度参数 (temperature parameter)** 来控制输出的平滑度。温度越高, 输出的概率分布越平滑 (即更接近均匀分布); 温度越低, 概率分布越尖锐。学生模型在较高温度下学习, 可以更好地理解类别之间的相对关系。

预训练模型)。然后在有限的数据集上进行剪枝和微调, 这样可以有效利用已有的知识, 减少对数据量的依赖。

- **动态更新标签:** 在有限数据集上进行微调时, 可以使用较小的学习率, 避免模型在小数据集上发生剧烈的参数调整, 从而防止过拟合。

(3) 标签精炼

- **噪声标签处理:** 在有限数据集上, 尤其是在数据扩展或知识蒸馏的过程中, 可能会引入一些噪声标签。标签精炼策略通过迭代地更新这些标签, 使得标签更加准确, 从而提高模型的性能。
- **动态更新标签:** 在微调过程中, 使用模型的预测来逐步调整和改进训练数据中的标签, 使其更接近实际分布。

(4) 结合预训练模型

- **迁移学习:** 在有限数据集上进行剪枝时, 通常会从一个大型数据集上预训练的模型开始 (即使用