深度学习——GAN 🖂

来自【机器学习面试题汇总与解析(蒋豆芽面试题总结)】 54 浏览 0 回复 2021-06-26



蒋豆芽 🕘



机器学习面试题汇总与解析——GAN

- 1. GAN是用来干什么的,怎么用的,介绍一下 \diamondsuit \diamondsuit \diamondsuit \diamondsuit
- 2. **GANs的优缺点是什么?** ☆ ☆ ☆ ☆
- 3. **GAN为什么不好收敛** ☆ ☆ ☆ ☆
- 4. **为什么GAN中的优化器不常用SGD** ☆ ☆ ☆ ☆ ☆
- 5. 生成对抗网络在哪里用到的,起什么作用,损失函数是什么 ☆ ☆ ☆ ☆ ☆
- 6. **训练GAN的一些技巧**☆ ☆ ☆ ☆ ☆
- 7. **说说GAN的训练过程** ☆ ☆ ☆ ☆
- 8. Pix2pix和cycleGan的区别 ☆ ☆ ☆ ☆ ☆

- 本专栏适合于Python已经入门的学生或人士,有一定的编程基础。
- 本专栏适合于**算法工程师、机器学习、图像处理求职**的学生或人士。
- 本专栏针对面试题答案进行了**优化,尽量做到好记、言简意赅。这才是一份面试题总结的正确打开** 方式。**这样才方便背诵**
- 如专栏内容有错漏,欢迎在评论区指出或私聊我更改,一起学习,共同进步。
- 相信大家都有着高尚的灵魂,请尊重我的知识产权,未经允许严禁各类机构和个人转载、传阅本专栏的内容。

关于**机器学习算法**书籍,我强烈推荐一本**《百面机器学习算法工程师带你面试》**,这个就很类似面经,还有讲解,写得比较好。私聊我进群。

关于**深度学习算法**书籍,我强烈推荐一本**《解析神经网络——深度学习实践手册》**,简称CNN book,通俗易懂。私聊我进群。

参考资料

B站机器学习视频: https://space.bilibili.com/10781175/channel/detail?cid=133301

:■ 蔣豆芽

pix2pix和CycleGAN: https://zhuanlan.zhihu.com/p/93219297?from_voters_page=true 读者可以把参考文章看看

1. **GAN是用来干什么的,怎么用的,介绍一下** \diamondsuit \diamondsuit \diamondsuit

参考回答

GAN的主要灵感来源于博弈论中零和博弈的思想,应用到深度学习神经网络上来说,就是通过生成网络G (Generator) 和判别网络D (Discriminator) 不断博弈,进而使G学习到数据的分布。

最初GAN被应用到图片生成上,则训练完成后,G可以从一段随机数中生成逼真的图像。后来 GAN的应用场景十分广泛,如图像生成、数据增强、图像编辑、恶意攻击检测、注意力预测、三 维结构生成等。

答案解析

- G, D的主要功能是:
- G是一个生成式的网络,它接收一个随机的噪声z(随机数),通过这个噪声生成图像
- D是一个判别网络,判别一张图片是不是"真实的"。它的输入参数是x,x代表一张图片,输出D(x)代表x为真实图片的概率,如果为1,就代表100%是真实的图片,而输出为0,就代表不可能是真实的图片

训练过程中,生成网络G的目标就是尽量生成真实的图片去欺骗判别网络D。而D的目标就是尽量辨别出G生成的假图像和真实的图像。这样,G和D构成了一个动态的"**博弈过程**",最终的平衡点即**纳什均衡点**

类似的问题还有:

2. **GAN的优缺点是什么?** ☆ ☆ ☆ ☆

参考回答

GAN的优缺点

优点:

- 1. **GANs**是一种以**半监督**方式训练分类器的方法,在没有很多带标签的训练集的时候,可以不做任何修改的直接使用GANs的代码
- 2. GAN是一种**生成式模型**,相比较其他生成模型(玻尔兹曼机和GSNs)只用到了**反向传播**,而不需要复杂的**马尔科夫链**
- 3. 相比其他所有模型, GAN可以产生更加清晰, 真实的样本

5. **GAN扩展到其他场景中十分简单**,比如图片风格迁移,超分辨率,图像补全,去噪,避免了损失函数设计的困难,不管三七二十一,只要有一个的基准,直接上**判别器**,剩下的就交给对抗训练

缺点:

- 1. 训练**GAN**需要达到**纳什均衡**,有时候可以用**梯度下降法**做到,有时候做不到。我们还没有找到很好的达到纳什均衡的方法,所以训练GAN相比VAE或者PixelRNN是不稳定的
- 2. GAN不适合处理离散形式的数据,比如文本
- 3. GAN训练不稳定。

答案解析

无。

3. **GAN为什么不好收敛** ☆ ☆ ☆ ☆ ☆

参考回答

- 1. 很难使一对模型 (G和D同时) 收敛。
- 2. 生成器G发生模式崩溃 (mode collapse, 即所有的输入图像生成相同的输出)
- 3. 生成器梯度消失问题

答案解析

生成对抗神经网络训练不稳定有如下三点原因:

- 1. 很难使一对模型(G和D同时)收敛,这也会造成以下第(2)点提到的模式崩溃。大多深度模型的训练都使用优化算法寻找损失函数比较低的值。优化算法通常是个可靠的"下山"过程。生成对抗神经网络要求双方在博弈的过程中达到**势均力敌(均衡)**。每个模型在更新的过程中(比如生成器)成功的"下山",但同样的更新可能会造成博弈的另一个模型(比如判别器)"上山"。甚至有时候博弈双方虽然最终达到了均衡,但双方在不断的抵消对方的进步并没有使双方同时达到一个有用的地方。对所有模型同时梯度下降使得某些模型收敛但不是所有模型都达到收敛最优。
- 2. 生成器G发生模式崩溃(mode collapse,即所有的输入图像生成相同的输出):对于不同的输入生成相似的样本,最坏的情况仅生成一个单独的样本,判别器的学习会拒绝这些相似甚至相同的单一样本。在实际应用中,完全的模式崩溃很少,局部的模式崩溃很常见。局部模式崩溃是指生成器使不同的图片包含相同的颜色或者纹理主题,或者不同的图片包含同一只狗的不同部分。MinBatch GAN缓解了模式崩溃的问题但同时也引发了counting, perspective和全局结构等问题,这些问题通过设计更好的模型框架有可能解决。

:■ 蔣豆芽

布,与真实数据分布相差距离太远,两个分布之间几乎没有任何重叠的部分,这时候判别器能够很快的学习把真实数据和生成的假数据区分开来达到判别器的最优,造成**生成器**的梯度无法继续更新甚至梯度消失。

4. **为什么GAN中的优化器不常用SGD** ☆ ☆ ☆ ☆

参考回答

- 1. **SGD容易震荡**,容易使GAN训练不稳定,
- 2. GAN的目的是在高维非凸的参数空间中找到**纳什均衡点**,GAN的纳什均衡点是一个鞍点,但是SGD只会找到**局部极小值**,因为SGD解决的是一个寻找最小值的问题,GAN是一个博弈问题。

答案解析

无。

5. 生成对抗网络在哪里用到的,起什么作用,损失函数是什么 \diamondsuit \diamondsuit \diamondsuit \diamondsuit

参考回答

GAN的主要灵感来源于博弈论中零和博弈的思想,应用到深度学习神经网络上来说,就是通过生成网络G (Generator) 和判别网络D (Discriminator) 不断博弈,进而使G学习到数据的分布。

最初GAN被应用到图片生成上,则训练完成后,G可以从一段随机数中生成逼真的图像。后来 GAN的应用场景十分广泛,如图像生成、数据增强、图像编辑、恶意攻击检测、注意力预测、三 维结构生成等。

GAN的**损失函数**是**JS散度**

$$\min_{G} \max_{D} \mathbb{E}_{\mathbf{x} \sim p_{data}}[\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}}[\log(1 - D(G(\mathbf{z})))].$$

答案解析

一切损失计算都是在**D (判别器)**输出处产生的,而D的输出一般是fake/true的判断,所以整体上采用的是**二进制交叉熵函数**。

左边包含两部分minG和maxD。

首先看一下maxD部分,因为训练一般是先保持G(生成器)不变训练D的。D的训练目标是**正确区分fake/true**,如果我们以1/0代表true/fake,则对第一项E因为输入采样自真实数据所以我们期望**D(x)趋近于1**,也就是第一项更大。同理第二项E输入采样自G生成数据,所以我们期望D(G(z))

第二部分保持D不变,训练G,这个时候只有第二项E有用了,关键来了,因为我们要迷惑D,所以这时将label设置为1(我们知道是fake,所以才叫迷惑),希望D(G(z))输出接近于1更好,也就是这一项越小越好,这就是minG。当然判别器哪有这么好糊弄,所以这个时候判别器就会产生比较大的误差,误差会更新G,那么G就会变得更好了

大概就是这样一个博弈过程了。

6. **训练GAN的一些技巧**☆ ☆ ☆ ☆ ☆

参考回答

- 1. 输入规范化到 (-1, 1) 之间, 最后一层的激活函数使用tanh (BEGAN除外)
- 2. 使用wassertein GAN的损失函数,
- 3. 如果有标签数据的话,尽量使用标签,也有人提出使用反转标签效果很好,另外使用标签平滑,单边标签平滑或者双边标签平滑
- 4. 使用mini-batch norm, 如果不用batch norm 可以使用instance norm 或者weight norm
- 5. 避免使用RELU和pooling层,减少稀疏梯度的可能性,可以使用leakrelu激活函数
- 6. 优化器尽量选择ADAM,学习率不要设置太大,初始1e-4可以参考,另外可以随着训练进行不断缩小学习率,
- 7. 给D的网络层增加高斯噪声, 相当于是一种正则

答案解析

无。

7. **说说GAN的训练过程** ☆ ☆ ☆ ☆

参考回答

略。

答案解析

无。

8. Pix2pix和cycleGan的区别 ☆ ☆ ☆ ☆ ☆

参考回答

1. pix2pix模型和CycleGAN都可以用于图像转换。

loss (cycle consistency loss)

4. pix2pix模型效果会更好; CycleGAN更容易训练。

答案解析

pix2pix模型,其贡献点在于提出了用GAN来解决图像转换问题的通用方法,并且证明了其方法的有效性。

pix2pix模型虽然可以取得较好的效果,但是它的训练需要成对的数据,而对于很多图像转换问题,成对数据是很难获取的甚至不可能。

CycleGAN背后的关键思想是它们可以建立在PIX2PIX架构的强大功能之上,但允许将模型指向两个离散的,不成对的图像集合。

CycleGAN的创新点在于**训练loss**,而基本上与无网络结构关系不大。但是相比pix2pix模型,CycleGAN训练难度依然很大。具体到训练细节上,CycleGAN的对抗loss采用**最小方差损失**,而且训练判别器时采用一些历史生成数据。



相关专栏



机器学习面试题汇总与解析(蒋豆芽面试题总结) 27篇文章 90订阅

已订阅

0条评论

○↑ 默认排序 ~



没有回复

请留下你的观点吧~

/ 牛客博客,记录你的成长

关于博客 意见反馈 免责声明 牛客网首页