

对于开发者来讲，2018年人工智能领域有哪些发展趋势

大家好，又见面了！在上一篇文章中，我谈到了自己关于研究领域的看法，即哪些研究领域正在成熟且能在今年发展壮大。继续从事研究当然很棒，但是，必定还有一些人工智能领域在2017年已经成熟、现已准备投入大规模应用了。这就是本文的主题——我想分享的是那些已经发展得足够好的技术。它们已能应用于你当下的工作中，甚至你能借此创业。重要提示：这是一份涵盖人工智能领域、算法和技术的清单，且它们都能立即投入使用。例如，你可以在文中看到时间序列分析，因为深度学习正在迅速取代信号处理领域之前的先进技术。但是，文中并没有提到强化学习（虽然它听起来更酷），因为在我看来它目前还不能投入工业应用。不过强化学习是个很了不起的、正在成长的研究领域。

此外，我想提醒你，这是一个包含三篇文章的系列文章之一，其中分别从三个角度分享了我对于明年人工智能领域会发生什么的想法：

作为正在推进领域发展的机器学习研究者

作为将机器学习应用于业界的机器学习开发人员（本文）

作为生活在这个新世界中的普通人

希望你能挑选适合自己的文章并享受阅读！

P.S.我在此并不会谈论图像识别和简单的计算机视觉，你们已经在这些领域努力多年啦。

GAN和虚假产物

即使几年前就创造出了生成式对抗网络（GAN），我对此一度非常怀疑。几年过去了，即使我看到GAN在生成64x64图像方面的巨大进步，我仍保持着怀疑。后来我阅读了一些数学方面的文章，文中提到GAN并不真正学习分布，我的疑虑愈发显著。不过今年发生了一些改变。首先，有趣的新型结构（例如CycleGAN）和数学改进（WassersteinGAN()）吸引我在实践中尝试了GAN，它们或多或少都工作得不错。在接下来的两个应用中，我彻底改变了态度，确信我们可以且必须利用它们来生成事物。

首先，我非常喜欢英伟达关于生成全高清图像的研究论文，而且它们看上去的确很真实（相比于一年前64x64的诡异面孔而言）：

AllofthesefacesarefakecelebritiesspawndbyAI()

不过，我最喜欢（作为一个完美的少年梦想应用）且被深深震撼的是生成虚假色情片：

AI-AssistedFakePornIsHereandWe'reAllFucked()

我还注意到许多游戏行业的应用，比如利用GAN生成景观、游戏主角甚至是整个世界。此外，我觉得我们必须对新的造假水平引起注意——包括你亲友的虚假色情片和线上完全虚假的个人。（可能不久之后线下也有？）

全体神经网络的独有模式

现代发展的问题之一（不仅局限于人工智能产业）是，我们有几十个不同的框架来完成同样的东西。今天。每个做机器学习的大公司都必然有自己的框架：谷歌、Facebook、亚马逊、微软、英特尔、甚至是索尼和Uber。此外，还有很多开源的解决方案！在一个简单的人工智能应用中，我们也希望使用不同的框架：例如，计算机视觉通常用Caffe2、自然语言处理常用PyTorch、推荐系统常用Tensorflow或Keras。把这些框架全部合并起来需要耗费大量的开发时间，并且会让数据科学家和软件开发者无法集中注意力完成更重要的任务。

解决方案必须是一个独一无二的神经网络形式。它需要能从任何框架中容易地获得，必须由开发人员轻松部署，并能让科学家轻易地使用。在这个问题上，今年出现了ONNX：

ONNX:OpenNeuralNetworkExchangeFormat()

实际上，它只是非循环计算图的简单格式，但却在实际中给了我们部署复杂人工智能解决方案的机会，而且我个人认为它非常具有吸引力——人们可以在像PyTorch这样的框架中开发神经网络，无需强大的部署工具，也不需要依赖Tensorflow的生态系统。

各类Zoo激增

对我来说，三年前人工智能界最让人兴奋的东西是CaffeZoo。当时，我正在做计算机视觉的相关工作，试遍所有模型，并检查它们如何工作、结果如何。稍后，我将这些模型应用于迁移学习或特征提取器。最近，我使用了两种不同的开源模型，类似大型计算机视觉流程中的一部分。这意味着什么？这意味着事实上没有必要训练自己的网络，例如ImageNet的对象识别或地点识别。这些基础的东西可以下载

并插入到你的系统当中。除了CaffeZoo之外，其他框架也有类似的Zoo。不过，让我感到惊奇的是，你可以仅在你的iPhone中插入计算机视觉、自然语言处理甚至加速度计信号处理的模型：

[likedan/Awesome-CoreML-Models\(\)](#)

我认为，这些Zoo只会越来越多，将ONNX这类生态系统的出现考虑在其中，并进行集中化（也会由于机器学习区块链应用而导致分散化）

自动机器学习替代流程

设计神经网络结构是件痛苦的任务——有时候你可以通过添加卷积层获得不错的结果；但是大多数时候，你需要使用超参数搜索方法（如随机搜索或贝叶斯优化）或是直觉仔细设计结构的宽度、深度和超参数。在计算机视觉领域，你至少可以调整在ImageNet上训练的DenseNet。但如果你在某些3D数据分类或多变量时间序列应用中工作，这一点将尤其困难。

使用其他神经网络生成神经网络从头生成神经网络结构的尝试有很多，但对我而言最棒、最清晰的是近期GoogleResearch的进展：

[AutoMLforlargescaleimageclassificationandobjectdetection\(\)](#)

他们使用AutoML生成的计算机视觉模型，比人类设计的网络工作的更快、更好！我相信，很快就会有許多关于这个话题的论文和开源代码。我认为，会出现更多博文和初创公司，告诉我们「我们人工智能所创造的人工智能学习了其他人工智能，它可以.....」，而不是「我们开发了一个人工智能，它可以.....」。至少，在我的项目中我会这样做。我也相信不是只有我一个人这样做。

智能堆栈正式化

在这个概念上，我阅读了很多AnatolyLevenchuk的博文。他是俄罗斯系统分析师、教练，同时热衷于人工智能领域。在下图中，你可以看到所谓「人工智能堆栈」的实例：

它不仅包括机器学习算法和你喜爱的框架，还有着更深的层次。而且在每个层面，都含有自身的发展和研究。

我认为，人工智能发展产业已经足够成熟，从而能拥有更多不同的专家。在团队中，仅有一名数据科学家远远不够——你需要不同人员从事硬件优化、神经网络研究

、人工智能编译、解决方案优化和生产实施等方面的工作。在他们之上必须有不同的团队领导、软件架构师（为每上述每个问题分别设计堆栈）以及管理人员。我曾经提过这个概念，希望将来人工智能领域的技术专家能够不断成长（对那些想成为人工智能软件架构师或技术引领者的人而言——你需要知道该学什么）。

基于语音的应用

人工智能所能解决的准确率可达95%以上的问题其实非常少：我们可以将图像识别分类到1000个类别，可以判断文本的正负面性、当然也能做一些更复杂的事情。我认为，还有一个领域将因人工智能派生的上千个应用发生动荡：那就是语音识别和生成。事实上，一年前DeepMind发布WaveNet之后这个领域还发展得不错。但是今天，多亏了百度的DeepVoice3和最近在GoogleTacotron2的发展，我们已经走远：

Tacotron2:GeneratingHuman-likeSpeechfromText()

这个技术很快就会发布在开源代码中（或着被一些聪明的人复制），而且每个人都能以非常高的准确率识别语音并生成它。这会带来什么呢？更好的私人助理、自动阅读器、谈判转录机.....当然还有作虚假声音产品。

更加智能化的机器人

我们今天所看到的机器人都有一个很大的问题——其中99%根本不基于人工智能，只是硬编码而已。因为我们意识到，我们不能用上百万个对话的注意力训练某些编码-译码器LSTM，从而获得智能系统。这就是为什么FacebookMessenger或Telegram中的绝大部分机器人都只有硬编码命令，至多拥有一些基于LSTM和word2vec的句子分类神经网络。但是，现代先进的NLP有点超出了这个水平。只需要看看Salesforce做出的有趣研究就能领会这一点：

AIResearch-Salesforce.com()

他们正在构建与数据库相连的NLP接口，克服现代编码-译码器的自动回归模型，不仅对文字或句子进行嵌入训练，而将范围扩展到了字符嵌入训练。此外，还有一个利用强化学习进行ROUGE分数NLP优化的有趣研究：

我相信，随着这些发展，我们至少能提升机器人水平。它们能检索更多智能信息、进行命名实体识别，而且很可能在某些封闭领域充分深度学习驱动机器人。

先进的序列分析

除Salesforce外，第二个被低估的公共机器学习研究实验室是Uber的人工智能实验室。前段时间，他们发表了一篇博客，展示了对时间序列的预测方法。老实说这让我有点受宠若惊，因为在我的应用中基本上也用了同样的方法！看看吧，这就是将统计特征和深度学习表示相结合的实例：

EngineeringExtremeEventForecastingatUberwithRecurrentNeuralNetworks()

如果你需要更多振奋人心的实例，请用34层一维ResNet尝试诊断心律失常。最棒的部分无疑是其性能——它不仅优于一些统计模型，甚至超过了专业心脏病专家的诊断！

Algorithmdiagnosesheartarrhythmiaswithcardiologist-levelaccuracy|StanfordNews()

最近，我从事的大部分就是深度学习的时间序列分析。我可以亲自证实神经网络工作得非常好，你能获得优于「黄金标准」5-10倍的性能。它真的能行！

超越内置的优化

我们如何训练神经网络？说实话，我们大多数人只是使用类似「Adam()」函数或是标准学习率。一些聪明的人选择最合适的优化器，并调整和调度学习率。我们总是低估「优化」这一主题，因为我们只需按下「优化」按钮，然后等待网络收敛就大功告成了。但是，在这个计算能力、存储和开源方案都大多相同的情况下，优胜者往往使用着与你相同的亚马逊实例，却能在Tensorflow模型中用最短的时间得到最佳的性能——这一切都是优化的功劳。

我鼓励大家看看上面Sebastian'sRuder的博文，其中谈到了2017年最新的标准优化器改进方案，以及其他非常有用的改善方法。你可以立即将其运用起来。

大肆宣传的整体下降

在阅读前文后，你能从这张图片中发现什么？考虑到许多开源工具和算法的发布，开发有价值的新东西并从中获取很大利润并不容易。我认为，对类似Prisma这样的初创公司而言，2018年并非最好的一年——将会有太多的竞争对手和「聪明人」想分一杯羹。他们可以将如今的开源网络部署在移动应用程序中，并称其为「创业」。

今年，我们必须专注于基础的事情，而不是很快获利——即使对于某些有声书初

创公司而言，我们计划用Google的Ratacon进行语音识别，但这也并非一个简单的网络服务，而是与合作伙伴携手的商业模式，同时也是吸引投资的商业模式。

总结

简而言之，有几种技术已经可以被用于实际产品：时间序列分析、GAN、语音识别、自然语言处理领域的部分发展。我们不应该再设计基础的分类或回归架构，因为AutoML会帮我们做这个。通过一些优化改进，我希望AutoML能够比以前运行得更快。此外，使用ONNX和各种模型Zoo能让我们仅用几行代码将基本模型注入到应用程序当中。我认为，制作人工智能为基础的应用，在目前先进的技术水平来说不是难事，而且对整个行业而言并无坏处！对于研究领域的发展，可以查阅我以前的文章。不久后，我将发布「2018人工智能发展趋势」系列的最后一篇文章，其中将介绍人工智能如何影响「普通人」的生活。

有没有可能实现计算机人工智能自主研究算法编译代码

现在对于一些需求简单的，有一些自动编程的工具，但是并非基于人工智能的。

基于人工智能的自动编程现在没有（至少我不知道有），而且很难。以后也许会出现，但是十年内应该不会有成熟的技术（更长的时间尺度就不好预测了。）

那么难点在哪呢？

首先，让计算机准确理解你的需求就很难很难。不要说计算机，即使是两个人之间，比如程序员和产品经理，程序员也通常并不能一次性的准确理解产品经理的需求。何况计算机？

其次，对于自动编程，要比下围棋难得多，因为围棋的目标是单一的，而编程要实现的功能则是各种各样。

编译原理是人工智能吗

编译原理是人工智能。人工智能（Artificial Intelligence），简称AI。它是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。人工智能是计算机科学的一个分支，它企图了解智能的实质，并生产出一种新的能以人类智能相似的方式做出反应的智能机器，该领域的研究包括机器人、语言识别、图像识别、自然语言处理和专家系统等。人工智能是对人的意识、思维的信息过程的模拟