**大数据与人工智能大致技术路线与预期效果**

深度学习，从大约1940开始就作为机器学习领域的一种算法被提出，但受阻于当时的计算能力以及数据量的不足这种算法一直都没有得到重视，直到最近几年得益于计算机技术的进步这种算法屡屡打破记录，尤其在 Alpha Go 击败 李世石 以后才真正崛起。  
  
所以我们真正想要学习的其实是机器学习中的深度学习算法。

1. 快速学习编程  
     
   编程是基础，但并不是说一定要学到可以找份工作才能学习深度学习。所以在开始学习之前用最快的速度掌握基本的Python语法，达到能够写简单的程序的水平。当然除了Python本身，还需要花点时间学习Numpy、Pandas以及Matplotlib这类数据科学库的使用方法。
2. 直接上手fast.ai  
     
   直接学习fast.ai的两门课程，能够快速地从应用的角度理解Deep Learning的技术原理。先从整体上了解这项技术的功能，再深入理解技术内的原理是个不错的方法。

fast.ai: Introduction to Machine Learning for Coders -- Jeremy Howard  
  
长久以来我们接受的都是传统的自下为上，也就是从基础知识到高级概念再到实践的教育。而在fast.ai的课程里，Jeremy Howard 提倡的是自上而下的学习方法，也就是回归到人最自然的学习方法。想想我们小时候是如何学习游戏、踢球、画画的，我们都是先按照自己的理解从模仿开始，随着对这样东西越来越熟悉再逐渐深入了解，学习如何把游戏玩得更好，研究踢球动作技巧，研究画画的笔法等等。在课程中，Jeremy 就是贯彻着这样的教学理念，从一开始就让大家先别管具体的原理，按照他所教的方法把机器学习的模型实现出来，让大家以最快的速度感受到这项技术的力量以及降低对它的恐惧感。  
  
这门课主要介绍了 Jeremy 自己最常用的两种机器学习的方法，一种是针对规则数据的Random Forest，另一种就是针对大部分非规则数据的Deep Learning。虽然没有吴恩达教授讲解得那么全面，但深入浅出，让学习者很容易理解这项技术的主要作用和原理。  
  
但这门课有一个门槛就是需要学习者最好有一年以上的编程经验，而且最好熟悉Python。  
  
  
fast.ai: Practical Deep Learning for Coders -- Jeremy Howard  
  
这门课才是fast.ai真正的重点课程，因为在这门课中Jeremy还介绍了他们专门为深度学习开发的一个库，利用这个库可以快速实现建模、导入数据以及训练模型。跟之前的课程一样，他还是采用的自上而下的教学方法，用最快的方式教会你使用他们的库，然后引导你用同样的方法去解决身边的问题。这门课还专门有个论坛让学习者们分享学习心得的，其中就有不少人在上过几堂课以后已经可以使用fast.ai框架训练出各种各样的模型，有的能根据卫星图分辨出图中的城市是属于哪个国家的，有的能根据植物的照片分辨出该植物是否患病，还有的能理解一些稀有语种的文字并自行造句。而这些都只需要上过三四堂课就能学会，上手难度非常低，让你在很短时间内就能领会到所谓的AI到底是什么、但同样的，这门课也是建议最好有一年以上编程经验。  
  
  
学完以上的课程以后基本上就能理解大部分的技术原理，接下来除了继续练习技术以外，还能根据自己的需要去学习更深层次的东西。

3. 加深概念理解  
  
在对技术有个大致的理解后，如果想要进一步深入，可以学习吴恩达教授的Deep Learning，从基本概念开始深入学习。在哔哩哔哩上先观看机器学习：吴恩达（Coursera/b站） 学习一些算法的概念，再通过  
liuyubobobo（慕课网） 用Python把所有算法真正实现一遍

对算法有了初步了解后应该直接学习吴恩达的深度学习

Coursera: Deep Learning -- Andrew Ng  
  
与吴恩达教授的另一门课Machine Learning课程相比，这门课只讲授了Deep Learning的相关内容。尽管如此，知识量却一点也没有少，而且都是专业领域中比较新的，所以可以通过这门课来了解专业比较近期的发展状态。除此之外，这门课的作业都是用Python以及基于它的深度学习框架如TensorFlow来完成的，所以在对于技术的了解以及后续的实用性上都是非常不错的。虽然在已经上了一门Machine Learning后对这门课的内容还是比较好接受的，但对于完全的初学者来说还是有一定难度。  
  
4.框架的学习

深度学习现在的主流框架是TensorFlow以及pytorch，但近两年使用pytorch的人越来越多。所以通过七月褚则伟（b站） IBM AI（Coursera）学习pytorch框架是一个不错的选择

5. 进阶深入理解  
  
跟进行业最新动态，最好的办法就是读最新的有重要突破的文献，并且找最简单的方式实验技术的成果。比如BERT就是一个预训练模型，所谓“预训练”就是能够用他们训练出来的模型只作简单的“微调整”训练就可以直接用在多种语言任务上并且获得非常好的表现。所以利用预训练模型，我们自己就可以很轻松地训练出一个处理其他特殊任务的模型，而不需要花费大量资源和时间从头开始训练。

以上内容是我借助互联网查询后逐步筛选后感觉适合自己的道路，目前仅初步学习完快速编程阶段的内容，后续学习内容与目标会根据自身情况而更改。