

深度学习笔记

陈鸿峥

2020.03*

目录

1 简介

1

本课程主要选用Ian Goodfellow, Yoshua Bengio, Aaron Courville的《深度学习》(Deep Learning)一书。

1 简介

图 1反映了**深度学习**与其他几个常见概念之间的关系。传统的**机器学习**（如决策树、SVM、随机森林等）常需要人工提取特征，这一步经常涉及到**特征工程**(feature engineering)，如果特征没有进行一定处理，直接丢进去让其学习，往往会产生非常糟糕的结果。在一种表示下可能可以对数据进行线性二分，而另一种表示下则没有办法。因此，为了避免对特征的强依赖性，一种方法是利用机器学习来学习**表示(representation)本身**，再将新的表示送入到后面的学习器中让它学习**表示到输出的映射**，此即**表示学习**。再到后来，深度学习则更加将这种思想发扬光大，表示学习只能学习到**浅层简单的特征**，那深度学习则尝试去学习**深层复杂的特征**。

事实上现在**图神经网络**(GNN)也是遵循这样的发展过程，最开始尝试在图上做机器学习[1, 2, 3]；然后又开始在图上以各种随机游走的方式做图表示学习-图嵌入(embedding)[4, 5]；后来发现图嵌入能够获得的特征依然太浅层了，因此现在更多则采用图神经网络[6, 7, 8, 9]的方式来做图相关的工作。

*Build 20200323

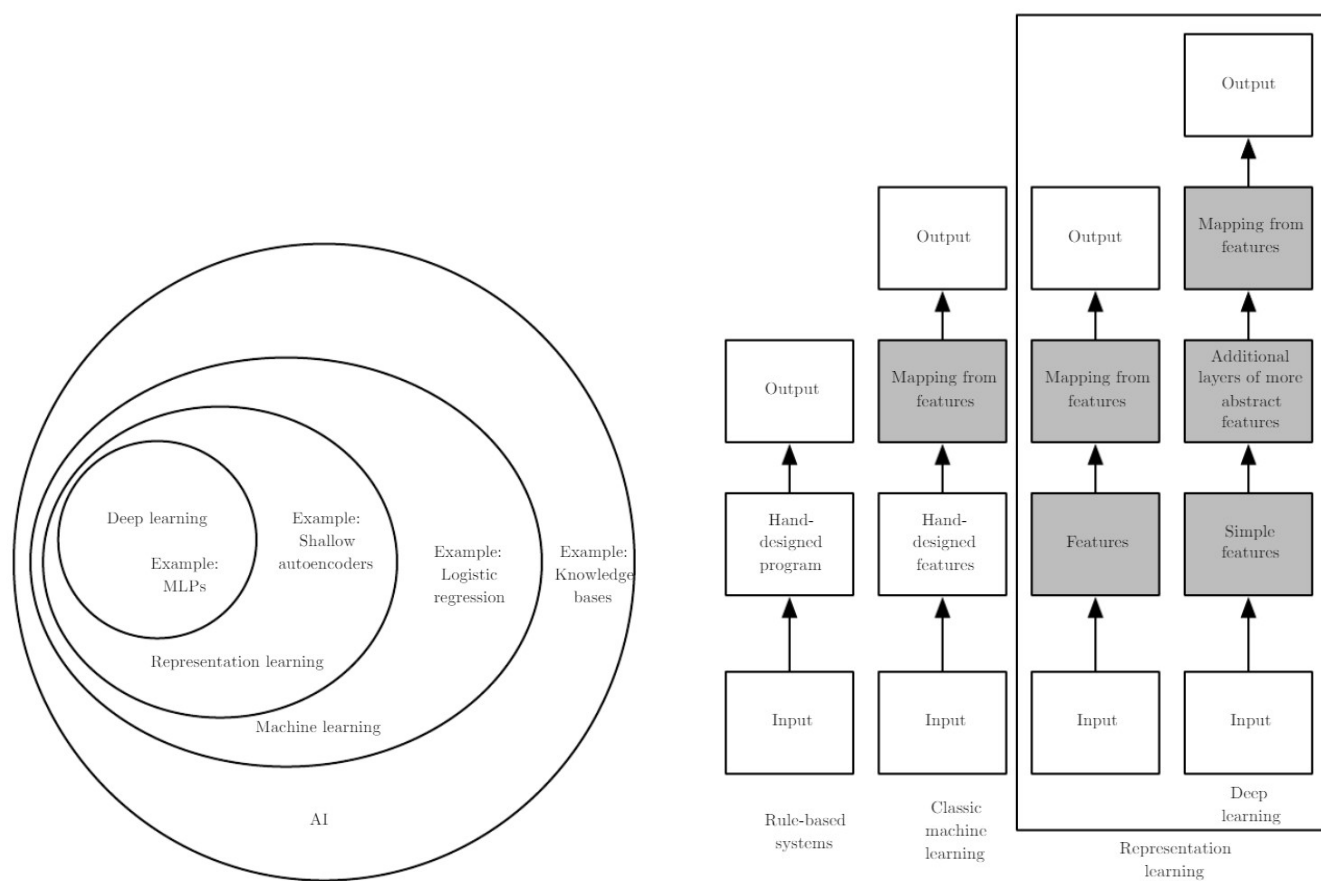


图 1: 深度学习Venn图

参考文献

- [1] Limin Yao, David Mimno, and Andrew McCallum. Efficient methods for topic model inference on streaming document collections. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2009.
- [2] Mu Li, David G Andersen, and Alexander Smola. Distributed delayed proximal gradient methods. In *NIPS Workshop on Optimization for Machine Learning*, 2013.
- [3] Rainer Gemulla, Erik Nijkamp, Peter J. Haas, and Yannis Sismanis. Large-scale matrix factorization with distributed stochastic gradient descent. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2011.
- [4] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. DeepWalk: Online learning of social representations. 2014.
- [5] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*, 2016.
- [6] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *Proceedings of the International Conference on Learning and Representation (ICLR)*, 2017.
- [7] William L. Hamilton, Rex Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *Advances in neural information processing systems (NeurIPS)*, 2017.
- [8] Yujia Li, Daniel Tarlow, Marc Brockschmidt, and Richard Zemel. Gated graph sequence neural networks. In *Proceedings of the International Conference on Learning and Representation (ICLR)*, 2016.
- [9] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *Proceedings of the International Conference on Learning and Representation (ICLR)*, 2018.