



中国科学院大学

University of Chinese Academy of Sciences

深度学习 (Deep Learning)

课程介绍

张新峰

计算机科学与技术学院

中国科学院大学

邮箱: xfzhang@ucas.ac.cn



计算机科学与技术学院

SCHOOL OF COMPUTER SCIENCE AND TECHNOLOGY

欢迎大家参加 《深度学习》课程！





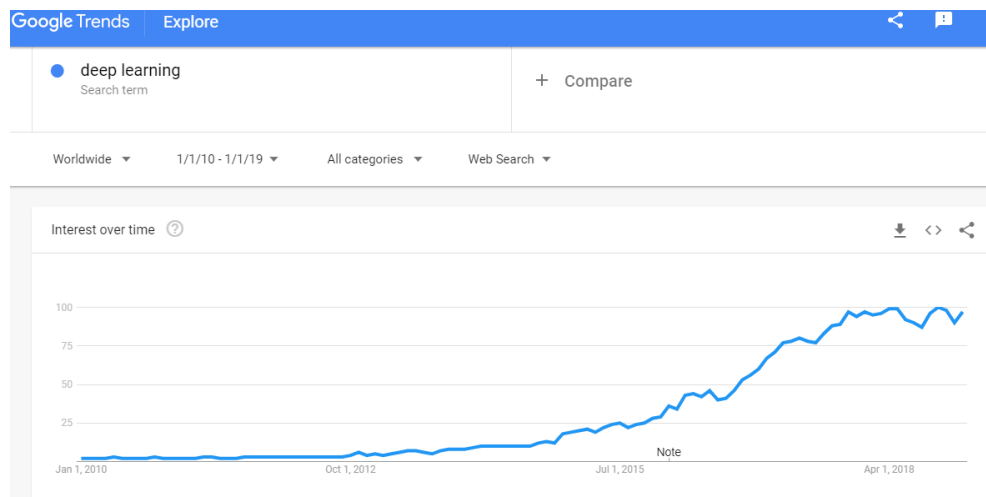
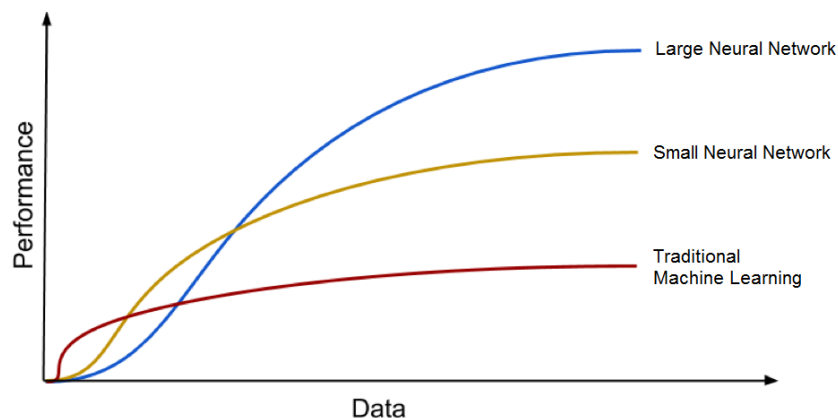
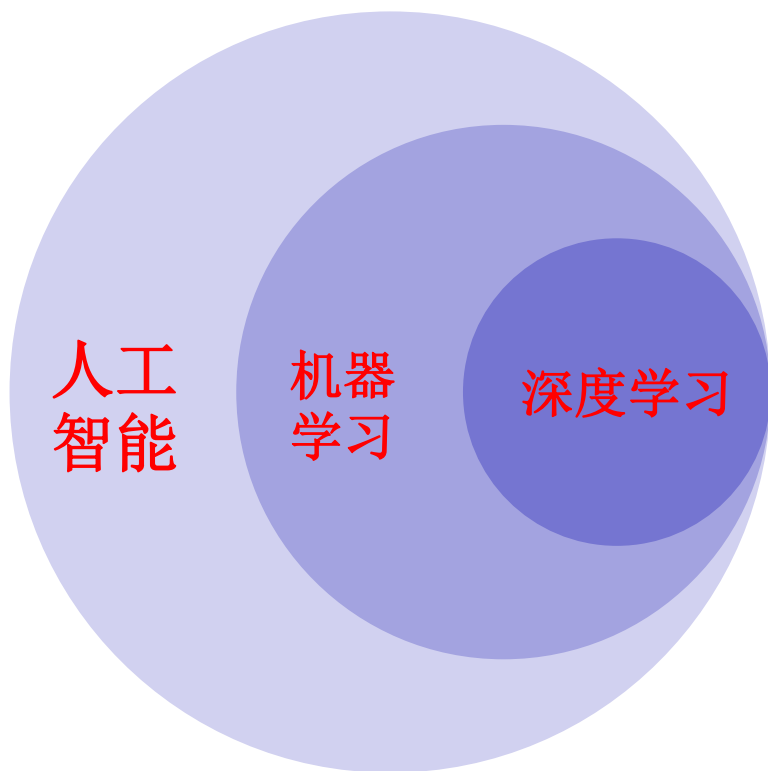
中国科学院大学
University of Chinese Academy of Sciences

提纲

- 课程信息
- 课程介绍
- 考核方式
- 参考资料

课程介绍

□ 深度学习是实现人工智能的一个工具或者技术手段



课程信息

□ 课程名称：《深度学习》

— 课程编号：081203M05009H

— 开课班次

- 深度学习1班

- 编号：081203M05009H-1

- 时间：每周一第9-11节课

- 授课教师：徐俊刚（教授），课程首席教师

- 深度学习2班

- 编号：081203M05009H-2

- 时间：每周二第9-11节课

- 授课教师：张新峰（副教授，长聘教轨助理教授）



课程信息

□ 深度学习2班授课团队介绍

— 主讲教师：张新峰

- 中科院计算所博士毕业，新加坡、美国、中国香港从事博士后研究工作，2019年加入国科大计算机学院，任新体制长聘教轨助理教授（副教授），博士生导师
- 研究方向：视频编码、质量评价、点云压缩和处理
- <https://dblp.org/pers/z/Zhang:Xinfeng.html>
- <https://scholar.google.com/citations?user=KQB-cKAAAAAJ&hl=en>
- 邮箱: xfzhang@ucas.ac.cn



课程信息

□ 深度学习2班授课团队介绍

— 教师助教

- 苏荔，国科大计算机学院，教授，suli@ucas.ac.cn
- 万方，国科大计算机学院，长聘教轨助理教授，wanfang@ucas.ac.cn

— 学生助教

- 李帅敏，国科大计算机学院，博士生，lishuaimin17@mails.ucas.ac.cn
- 洪铭遥，国科大计算机学院，博士生，hongmingyao14@mails.ucas.ac.cn
- 毕超，国科大计算机学院，博士生，bichao18@mails.ucas.ac.cn



课程介绍

❑ 课程类型：专业普及课

❑ 学时/学分：40/2

❑ 先修课程要求：机器学习，线性代数，概率论和统计，
C++/Python

– 常见平台：Pytorch, TensorFlow, Keras, Caffe,
PaddlePaddle

❑ 课程主要内容

- 讲授和讨论深度学习的主要理论和关键技术，主要内容有深度学习基础、卷积神经网络、循环神经网络、深度生成模型、深度学习正则化等以及上述深度学习理论在图像、语音、自然语言处理等领域的主要应用，同时也介绍了一些新兴的深度学习模型及其应用
- 课程注重深度学习实践能力的锻炼和培养，通过引入多个深度学习课程实验，提升同学们的动手能力



课程介绍

- ❑ 第一章 引言
- ❑ 第二章 深度学习基础
- ❑ 第三章 卷积神经网络
- ❑ 第四章 循环神经网络
- ❑ 第五章 深度生成模型
- ❑ 第六章 其他典型深度学习方法
- ❑ 第七章 深度学习中的正则化
- ❑ 第八章 深度学习工具
- ❑ 第九章 深度学习在图像识别中的典型应用
- ❑ 第十章 深度学习在语音识别中的典型应用
- ❑ 第十一章 深度学习在自然语言处理中的典型应用



考核方式

□ 考核内容, 方式和百分比

- 开卷笔试: **45%**
- 文献阅读: **10%**
 - 10篇相关论文, 提交中文版的PPT介绍
 - 每篇论文提交一个PPT (10页以上)

□ 课程作业: **45%**

- 视觉相关
 - 手写数字识别
 - 猫狗分类
- 自然语言处理相关
 - 自动写诗
 - 电影评论情感分类



考核方式

□ 选做部分的课程作业

- 附加题，可额外加分，可组队完成（每组不超过3人）
- 视觉相关
 - 车牌识别
 - 行人检测
 - 视频超分辨率
- 自然语言处理
 - 神经网络语言模型
 - 神经机器翻译



考核方式

□ 选做部分可参加竞赛

- 取得奖励的团队，其成员课程作业满分
- 取得有效成绩的团队，其成员课程作业适当加分

□ 2021年相关赛事

- 2021年中国人工智能大赛：<https://ai.xm.gov.cn/detail/intro-detail.html?id=2e41afb296d14c608ea8488084ef13b9>
- 全国人工智能大赛（每年举行）：
<https://naic.pcl.ac.cn/homepage/index.html>
- 百度大脑（持续）：<https://aistudio.baidu.com/aistudio/competition>
- 华为人工智能校园创新大赛（持续）：
<https://developer.huawei.com/consumer/cn/activity/digixActivity/digixdetail/101592535978459755>



参考资料(书籍)

- ❑ [1] Ian, Goodfellow等著, 赵申剑等译. 深度学习. 北京: 人民邮电出版社, 2017.
- ❑ [2] 山下隆义著, 张弥译. 图解深度学习. 北京: 人民邮电出版社, 2018.
- ❑ [3] Yoav Goldberg著, 车万翔等译. 基于深度学习的自然语言处理. 北京: 机械工业出版社, 2018.
- ❑ [4] 猿辅导研究团队. 深度学习核心技术与实践. 北京: 电子工业出版社, 2018.
- ❑ [5] 林大贵. TensorFlow+Keras深度学习人工智能实践应用. 北京: 清华大学出版社, 2018.
- ❑ [6] 刘祥龙等著. PaddlePaddle深度学习实战. 北京: 机械工业出版社, 2018.



参考资料(网站)

- ❑ <http://openclassroom.stanford.edu/MainFolder/CoursePage.php?course=MachineLearning>
- ❑ <http://deeplearning.net/tutorial/>
- ❑ <http://neuralnetworksanddeeplearning.com/>
- ❑ <http://pytorch123.com/#pytorch>



参考资料 (论文)

- [1] G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks, *Science*, 2006, 313 (5786): 504 -507.
- [2] G. E. Hinton, S. Osindero, Y. W. Teh. A fast learning algorithm for deep belief nets. *Neural computation*, 2006, 18(7): 1527-1554.
- [3] Y. LeCun and Y. Bengio. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 1995.
- [4] N. Roux Le, and Y. Bengio. Representational power of restricted Boltzmann machines and deep belief networks. *Neural Computation*, 2008, 20(6): 1631-1649.
- [5] J. Ngiam, A. Khosla, and M. Kim. Multimodal deep learning. *Proceedings of International Conference on Machine Learning*, 2011: 689-696.
- [6] A. Graves, A. Mohamed, and G. E. Hinton. Speech recognition with deep recurrent neural networks. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013.
- [7] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 1997.



参考资料 (论文)

- [8] S. Sukhbaatar, J. Weston, R. Fergus. End-to-end memory networks. *Advances in Neural Information Processing Systems*, 2015: 2431-2439.
- [9] D. Bahdanau, K. Cho, and Y. Bengio. Neural machine translation by jointly learning to align and translate. *Proceedings of the International Conference on Learning Representations*, 2015.
- [10] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair et al., and Y. Bengio. Generative adversarial nets. *Advances in Neural Information Processing Systems*, 2014.
- [11] S. Sabour, N. Frosst, G. E. Hinton. Dynamic Routing Between Capsules. *CoRR*, abs/1710.09829, 2017.
- [12] J. Xu, H. Li, S. Zhou. Improving mixing rate with tempered transition for learning restricted Boltzmann machines. *Neurocomputing*, 2014, 139:328-335.
- [13] V. Mnih, N. Heess, A. Graves, K. Kavukcuoglu. Recurrent models of visual attention. *Advances in Neural Information Processing Systems*, 2014: 2204-2212.
- [14] K. Xu, J. Ba, R. Kiros. Show, attend and tell: Neural image caption generation with visual attention. *Proceedings of the International Conference on Machine Learning*, 2015.



参考资料 (论文)

- [15] K. M. Hermann, T. Kocisky, E. Grefenstette. Teaching machines to read and comprehend. *Advances in Neural Information Processing Systems*, 2015: 1684-1692.
- [16] Y. Kim. Convolutional neural networks for sentence classification. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014: 1746-1751.
- [17] T. Mikolov, M. Karafiát, L. Burget, et al. Recurrent neural network based language model. *Proceedings of Eleventh Annual Conference of the International Speech Communication Association*, 2010: 1045–1048.
- [18] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014
- [19] R. Kiros, R. Salakhutdinov, and R. S. Zemel. Unifying visual semantic embeddings with multi-modal neural language models. *Transactions of the Association for Computational Linguistics*, 2015.
- [20] Z. Zhou, J. Feng. Deep Forest: Towards an alternative to deep neural networks. *Proceedings of International Joint Conference on Artificial Intelligence*, 2017.



参考资料 (论文)

- [21] J. Mao, W. Xu, Y. Yang, J. Wang, Z. Huang, and A. Yuille. Deep captioning with multimodal recurrent neural networks (m-rnn). Proceedings of the International Conference on Learning Representations, 2014.
- [22] J. Donahue, L. A. Hendricks. Long-term Recurrent Convolutional Networks for Visual Recognition and Description. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2015
- [23] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan. Show and tell: A neural image caption generator. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2015.
- [24] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2016
- [25] G. Jiaxian, L. Sidi, C. Han, Z. Weinan, W. Jun, Y. YuLong. Text generation via adversarial training with leaked information. In Thirty-Second AAAI Conference on Artificial Intelligence, 2018.



谢谢！

