November 12, 2020

Introduction

Thanks for the feedback. Every single feedback response requested more information on the topic hence here we are.

- Lack of blackboard for math compounded with remote presentation - feel free to ask questions. Ambitious agenda so I might not be able to get to all questions but will try.
- The group is comfortable with linear algebra, multivariate calculus, gradients and optimization in multiple dimensions etc. Very little linear algebra in the presentation but you will need it.
- What will not be covered - refer to the slide.
- Objective - you should feel like can start playing with python code now.

Convolution network

Discuss why architecture matters.

- Exploit relationships better
- Numerical stability/ease of computation.

ImageNet challenge

- Describe the challenge
- Start taking about what happened in 2012

Transfer learning

Motivation

- People not only build models but also make the trained weights available. We will see a little bit of pytorch code. Data, fine tuning and execution all reused.
- We may not have enough data to train a deep model.
- Our task can exploit features detected by an existing pretrained model.

Adversarial example

- Start with Rohan's work
- Introduce adversarial network
- Discuss stop sign, trucks with LCDs

References

[1] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," 2018.

[2] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, pp. 2278–2324, Nov 1998.

[3] I. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," in *International Conference on Learning Representations*, 2015.

[4] A. Kurakin, I. Goodfellow, and S. Bengio, "Adversarial examples in the physical world," *ICLR Workshop*, 2017.

[5] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning.* The MIT Press, 2016.

[6] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Advances in Neural Information Processing Systems 27* (Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, eds.), pp. 3104–3112, Curran Associates, Inc., 2014.

[7] S. J. Russell and P. Norvig, *Artificial Intelligence: a modern approach.* Pearson, 3 ed., 2009.