

# Deep Learning

## An MIT Press book in preparation

Ian Goodfellow, Yoshua Bengio and Aaron Courville

[Book](#) [Exercises](#) [External Links](#)

## Lectures

We plan to offer lecture slides accompanying all chapters of this book. We currently offer slides for only some chapters. If you are a course instructor and have your own lecture slides that are relevant, feel free to contact us if you would like to have your slides linked or mirrored from this site.

1. [Introduction](#)
  - Presentation of Chapter 1, based on figures from the book [[.key](#)] [[.pdf](#)]
  - [Video](#) of lecture by Ian and discussion of Chapter 1 at a reading group in San Francisco organized by Alena Kruchkova
2. [Linear Algebra](#) [[.key](#)] [[.pdf](#)]
3. [Probability and Information Theory](#) [[.key](#)] [[.pdf](#)]
4. [Numerical Computation](#) [[.key](#)] [[.pdf](#)] [[youtube](#)]
5. [Machine Learning Basics](#) [[.key](#)] [[.pdf](#)]
6. [Deep Feedforward Networks](#) [[.key](#)] [[.pdf](#)]
  - [Video](#) (.flv) of a presentation by Ian and a group discussion at a reading group at Google organized by Chintan Kaur.
7. [Regularization for Deep Learning](#) [[.pdf](#)] [[.key](#)]
8. [Optimization for Training Deep Models](#)
  - **Gradient Descent and Structure of Neural Network Cost Functions** [[.key](#)] [[.pdf](#)]  
 These slides describe how gradient descent behaves on different kinds of cost function surfaces. Intuition for the structure of the cost function can be built by examining a second-order Taylor series approximation of the cost function. This quadratic function can give rise to issues such as poor conditioning and saddle points. Visualization of neural network cost functions shows how these and some other geometric features of neural network cost functions affect the performance of gradient descent.
  - **Tutorial on Optimization for Deep Networks** [[.key](#)] [[.pdf](#)]  
 Ian's presentation at the 2016 Re-Work Deep Learning Summit. Covers Google Brain research on optimization, including visualization of neural network cost functions, Net2Net, and batch normalization.
  - **Batch Normalization** [[.key](#)] [[.pdf](#)]
  - **Video of lecture / discussion:** This video covers a presentation by Ian and group discussion on the end of Chapter 8 and entirety of Chapter 9 at a reading group in San Francisco organized by Taro-Shigenori Chiba.
9. [Convolutional Networks](#)
  - **Convolutional Networks** [[.key](#)] [[.pdf](#)]  
 A presentation summarizing Chapter 9, based directly on the textbook itself.
  - **Video of lecture / discussion:** This video covers a presentation by Ian and group discussion on the end of Chapter 8 and entirety of Chapter 9 at a reading group in San Francisco organized by Taro-Shigenori Chiba.
10. [Sequence Modeling: Recurrent and Recursive Networks](#)
  - **Sequence Modeling** [[.pdf](#)] [[.key](#)]  
 A presentation summarizing Chapter 10, based directly on the textbook itself.
  - **Video of lecture / discussion.** This video covers a presentation by Ian and a group discussion of Chapter 10 at a reading group in San Francisco organized by Alena Kruchkova.
11. [Practical Methodology](#) [[.key](#)] [[.pdf](#)] [[youtube](#)]
12. [Applications](#) [[.key](#)] [[.pdf](#)]
13. [Linear Factors](#) [[.key](#)] [[.pdf](#)]
14. [Autoencoders](#) [[.key](#)] [[.pdf](#)]
15. [Representation Learning](#) [[.key](#)] [[.pdf](#)]
16. [Structured Probabilistic Models for Deep Learning](#) [[.key](#)] [[.pdf](#)]
17. [Monte Carlo Methods](#) [[.key](#)] [[.pdf](#)]
18. [Confronting the Partition Function](#) [[.key](#)] [[.pdf](#)]