

# Lectures

Below is a summary of the lecture topics and links to the lecture slides. I will try and make all slides available before the lecture begins. We might vary the order of the lecture topics (probability of this happening is larger for the later lectures). The topics of Lectures 1-5 are fairly set though.



## Lecture 1

**Title:** *The Deep Learning Revolution*

**Date:** March 18

**Topics covered:**

- Review of the impact deep networks have had in the application fields of speech, computer vision and NLP.
- Review of the course's syllabus.
- Review of the course's assignments, project and assessment.

**Slides:** [Lecture1.pdf \(https://kth.instructure.com/courses/7507/files/1822869/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1822869/download?wrap=1)  <https://kth.instructure.com/courses/7507/files/1822869/download?wrap=1>, [Lecture1 CourseAdmin.pdf \(https://kth.instructure.com/courses/7507/files/1822868/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1822868/download?wrap=1)  <https://kth.instructure.com/courses/7507/files/1822868/download?wrap=1>

## Lecture 2

**Title:** *Learning Linear Binary & Linear Multi-class Classifiers from Labelled Training Data*

(mini-batch gradient descent optimization applied to "Loss + Regularization" cost functions)

**Date:** March 19

**Topics covered:**

- Supervised learning = minimizing loss + regularization.
- Learning linear classifiers.
- Binary SVM classifiers as an unconstrained optimization problem.
- Gradient Descent, SGD, mini-batch optimization.
- Multi-class classification with one layer networks.
- Different loss-functions.

**Slides:** [Lecture2.pdf \(https://kth.instructure.com/courses/7507/files/1830393/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1830393/download?wrap=1)  <https://kth.instructure.com/courses/7507/files/1830393/download?wrap=1>

**Suggested Reading Material:** Sections 5.1.4, 5.2, 5.2.2, 5.7.2 from "Deep Learning" by Goodfellow, Bengio and Courville. Link to [Chapter 5 of Deep Learning \(http://www.deeplearningbook.org/contents/ml.html\)](http://www.deeplearningbook.org/contents/ml.html)

Sections 8.1.3, 8.3.1 from the book give amore detailed description and analysis of mini-batch gradient descent and SGD than given in the lecture notes. Link to [Chapter 8 of Deep Learning](#)

<http://www.deeplearningbook.org/contents/optimization.html>).

The suggested readings from chapter 5 should be familiar to those who have already taken courses in ML. Lecture 2 should be more-or-less self-contained. But the reading material should flesh out some of the concepts referred to in the lecture.


## Lecture 3

**Title:** *Back Propagation*

**Date:** March 21

**Topics covered:**

- Chain rule of differentiation.
- Chain rule for vector inputs and outputs
- Computational graph.
- Back propagation (In more detail than you probably ever expected!)

**Slides:** [Lecture3.pdf \(https://kth.instructure.com/courses/7507/files/1830788/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1830788/download?wrap=1)   
<https://kth.instructure.com/courses/7507/files/1830788/download?wrap=1>

**Suggested Reading Material:**

Section 6.5 from the deep learning book.

I'm going to go into very explicit detail about the back-propagation algorithm. It was not my original intention to have such an involved description but condensing the explanation made things less clear. My hope, though, is that everybody will have a good understanding of the theory and the mechanics of the algorithm after this lecture. I go into more specific detail (but not as generic) than in the deep learning book. So my recommendation is that you read my lecture notes to get a good understanding for the concrete example(s) I explain and then you can read the deep learning book for a broader description.

Section 6.5 also assume you know about networks with more than 1 layer! So it may be better to hold off reading it until after lecture 4 (where I will talk about n-layer networks, activation functions, etc..)

## Lecture 4

**Title:** *k-layer Neural Networks*

**Date:** March 25

**Topics covered:**

- k-layer Neural Networks.
- Activation functions.
- Backprop for k-layer neural networks.
- Problem of vanishing and exploding gradients.
- Importance of careful initialization of network's weight parameters.
- Batch normalization + Backprop with Batch normalisation

**Slides:** [Lecture4.pdf \(https://kth.instructure.com/courses/7507/files/1841661/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1841661/download?wrap=1)   
<https://kth.instructure.com/courses/7507/files/1841661/download?wrap=1>

**Suggested Reading Material:**

Sections 8.7.1 from the deep learning book has a more subtle description of the benefits of batch normalisation and why it works.

## Lecture 5

**Title:** *Training & Regularization of Neural Networks*

**Date:** March 26

**Topics covered:**

- The art/science of training neural networks.
- Hyper-parameter optimisation.
- Variations of SGD.
- Regularization via DropOut.
- Evaluation of the models - ensembles

**Slides:** [Lecture5.pdf \(https://kth.instructure.com/courses/7507/files/1852571/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1852571/download?wrap=1)   
(<https://kth.instructure.com/courses/7507/files/1852571/download?wrap=1>)

**Suggested Reading Material:**

Sections 8.3.1, 8.3.2, 8.3.3, 8.5 from the deep learning book cover variations of the SGD in detail.

## Lecture 6

**Title:** *All about Convolutional Networks*

**Date:** Apr 1

**Topics covered:**

- Details of the convolution layer in Convolutional Networks.
- Gradient computations for a convolutional layers.
- Common operations in ConvNets - max-pooling etc

**Slides:** [Lecture6.pdf \(https://kth.instructure.com/courses/7507/files/1852573/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1852573/download?wrap=1)   
(<https://kth.instructure.com/courses/7507/files/1852573/download?wrap=1>)

**Suggested Reading Material:**

Section 9.1, 9.2 (motivates benefit of convolutional layers Vs fully connected layers), 9.10 (if you are interested in the neuro-scientific basis for ConvNets). Section 9.3 discusses the pooling operation.

## Lecture 7

**Title:** *Training & Designing ConvNets*

**Date:** Apr 2

**Topics covered:**

- Review of the modern top performing deep ConvNets - AlexNet, VGGNet, GoogLeNet, ResNet
- Practicalities of training deep neural networks - data augmentation, transfer learning and stacking convolutional filters.
- The Transposed Convolution operation
- Bottleneck ConvNets (Semantic Segmentation, Depth estimation)
- Attempts at Visualizing what a deep ConvNet has learnt

**Slides:** [Lecture7.pdf \(https://kth.instructure.com/courses/7507/files/1867007/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1867007/download?wrap=1)   
<https://kth.instructure.com/courses/7507/files/1867007/download?wrap=1>



## Lecture 8

**Title:** Deep Learning Frameworks and Computational Facilities

**Date:** Apr 8

**Topics covered:**

- General purpose software for deep learning Tensorflow & PyTorch
- Fundamental differences between Tensorflow & PyTorch - static and dynamic computation graphs
- How to use the Google Cloud service

**Slides:** [Lecture8.pdf \(https://kth.instructure.com/courses/7507/files/1864796/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1864796/download?wrap=1)   
<https://kth.instructure.com/courses/7507/files/1864796/download?wrap=1>, [Lecture8 Project admin.pdf \(https://kth.instructure.com/courses/7507/files/1864765/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1864765/download?wrap=1)   
<https://kth.instructure.com/courses/7507/files/1864765/download?wrap=1>

## Lecture 9

**Title:** *Networks for Sequential Data: RNNs & LSTMs*

**Date:** April 9

**Topics covered:**

- RNNs.
- Back-prop for RNNs.
- RNNs for synthesis problems.
- Problem of exploding and vanishing gradients in RNN.
- LSTMs
- ConvNets Vs LSTMs for sequential data

**Slides:** [Lecture9.pdf \(https://kth.instructure.com/courses/7507/files/1866442/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1866442/download?wrap=1)   
<https://kth.instructure.com/courses/7507/files/1866442/download?wrap=1>

## Lecture 10

**Title:** *Networks for Sequential Data: RNNs & LSTMs Ctd*

**Date:** April 23

**Topics covered:**

- Finished the content from the previous lecture
- Tutorial examples on the board

**Slides:**


## Lecture 11

**Title:** *Unsupervised learning of how to generate realistic image*

**Date:** April 25

**Topics covered:**

- History of (Deep) Generative Modeling
- Variational Auto-encoders
- Generative Adversarial Networks
- Adversarial Training
- PixelRNN

**Slides:** [Lecture 11.pdf \(https://kth.instructure.com/courses/7507/files/1889979/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1889979/download?wrap=1)   
[\(https://kth.instructure.com/courses/7507/files/1889979/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1889979/download?wrap=1)

## Lecture 12

**Title:** Deep learning for translation problems

**Date:** April 29

**Topics covered:**

- Deep Learning and NLP embeddings
- Language Translation
  - Supervised Learning
- Image Captioning
- Attention Networks

**Slides:** [Lecture12.pdf \(https://kth.instructure.com/courses/7507/files/1895873/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1895873/download?wrap=1)   
[\(https://kth.instructure.com/courses/7507/files/1895873/download?wrap=1\)](https://kth.instructure.com/courses/7507/files/1895873/download?wrap=1)

## Lecture 13

**Title:** *Transformer Networks & some odds and ends*

**Date:** 30th April

**Topics covered:**

- Transformer networks & self-attention
- Adversarial examples
- Some Odds & ends
- Q & A from students.

**Slides:**