Introduction to Deep Learning

Charles Ollion - Olivier Grisel
Goal of the class

Overview

- When and where to use DL
- "How" it works
- Frontiers of DL
Goal of the class

Overview

- When and where to use DL
- "How" it works
- Frontiers of DL

Arcanes of DL

- Implement using Numpy, and Keras
- Engineering knowledge for building and training DL
What is Deep Learning

Good old Neural Networks, with more layer/modules
What is Deep Learning

Good old Neural Networks, with more layer/modules

Non-linear, hierarchical, abstract representations of data
What is Deep Learning

Good old Neural Networks, with more layer/modules

Non-linear, hierarchical, abstract representations of data

Flexible models with any input/output type and size
What is Deep Learning

Good old Neural Networks, with more layer/modules

Non-linear, hierarchical, abstract representations of data

Flexible models with any input/output type and size

Differentiable Functional Programming
Typical ML system

Data independent

Feature Extractions

Classification

Supervised Learning

dog
Typical ML system
Deep Learning system
Why Deep Learning Now?

- Better algorithms & understanding
- Computing power (GPUs, TPUs, ...)
- Data with labels
- Open source tools and models
Why Deep Learning Now?

- Better algorithms & understanding
- Computing power (GPUs, TPUs, ...)
- Data with labels
- Open source tools and models
Why Deep Learning Now?

- Better algorithms & understanding
- Computing power (GPUs, TPUs, ...)
- Data with labels
- Open source tools and models
Why Deep Learning Now?

- Better algorithms & understanding
- Computing power (GPUs, TPUs, ...)
- Data with labels
- Open source tools and models
DL Today: Speech-to-Text

[Baidu 2014]
DL Today: Vision
DL Today: Vision
DL Today: NLP

[Google Translate System - 2016]

[Socher 2015]
DL Today: NLP

Salit Kulla

to me

Hey, Wynton Marsalis is playing this weekend. Do you have a preference between Saturday and Sunday?

-S

[Google Inbox Smart Reply]
DL Today: NLP

Hey, Wynton Marsalis is playing this weekend. Do you have a preference between Saturday and Sunday?

-S

Most of chatbots claiming "AI" do not use Deep Learning (yet?)
DL Today: Vision + NLP
DL Today: Image translation
DL Today: Generative models

Sampled celebrities [Nvidia 2017]
DL Today: Generative models

Sampled celebrities [Nvidia 2017]
DL Today: Generative models

Sound generation with WaveNet [DeepMind 2017]
DL Today: Generative models

Sound generation with WaveNet [DeepMind 2017]

Guess which one is generated?

Tacotron 2 Natural TTS Synthesis by Conditioning WaveNet on Mel Spectrogram
DL in other sciences

[Deep Genomics 2017]
DL in other sciences

[Deep Genomics 2017]
DL in other sciences

[Deep Genomics 2017]
DL for AI in games

[Deepmind AlphaGo / Zero 2017]

[Atari Games - DeepMind 2016]

[Starcraft 2 for AI research]
DL for AI in games

[Deepmind AlphaGo / Zero 2017]

[Atari Games - DeepMind 2016]

[Starcraft 2 for AI research]
Outline of the class

Backpropagation
Outline of the class

Backpropagation

Computer Vision (1 & 2)
Outline of the class

Backpropagation

Computer Vision (1 & 2)

Recommender Systems
Outline of the class

Backpropagation

Computer Vision (1 & 2)

Recommender Systems

Natural Language Processing (1 & 2)
Outline of the class

Backpropagation

Computer Vision (1 & 2)

Recommender Systems

Natural Language Processing (1 & 2)

Optimization: theory, methods and tricks (2)
Outline of the class

Backpropagation

Computer Vision (1 & 2)

Recommender Systems

Natural Language Processing (1 & 2)

Optimization: theory, methods and tricks (2)

Generative models and unsupervised learning
How this unit works

Lecture 1h-1h30

Coding sessions 2h-2h30

- 5 min multiple choice evaluation of previous lab
- split into 2 groups
- BYO laptop, work by pairs
- Homework 3h per week
How this unit works

Lecture 1h-1h30

Coding sessions 2h-2h30

- 5 min multiple choice evaluation of previous lab
- split into 2 groups
- BYO laptop, work by pairs
- Homework 3h per week

Final exam 2h
How this unit works

Lecture 1h-1h30

Coding sessions 2h-2h30

- 5 min multiple choice evaluation of previous lab
- split into 2 groups
- BYO laptop, work by pairs
- Homework 3h per week

Final exam 2h

Recommended reading: deeplearningbook.org, Francois Chollet's
Frameworks and Computation Graphs
 Libraries & Frameworks

Libraries & Frameworks

Automatic differentiation: **TensorFlow**, MXnet, CNTK, **Theano**

Dynamic and high level: Torch & **PyTorch**, Chainer, MinPy, DyNet...
Libraries & Frameworks

Automatic differentiation: **TensorFlow**, MXnet, CNTK, *Theano*

Dynamic and high level: Torch & **PyTorch**, Chainer, MinPy, DyNet...

**Keras**: high level frontend for TensorFlow, mxnet, theano, cntk
Computation Graph

Neural network = parametrized, non-linear function
Computation Graph

Computation graph: Directed graph of functions, depending on
Computation Graph

Combination of linear (parametrized) and non-linear functions
Computation Graph

Not only sequential application of functions
Computation Graph

Automatic computation of gradients: all modules are **differentiable**!
Computation Graph

Automatic computation of gradients: all modules are **differentiable**!

**Tensorflow, theano, etc. build a static computation graph**
Computation Graph

Automatic computation of gradients: all modules are differentiable!

Tensorflow, theano, etc. build a static computation graph

Torch, pytorch, etc. rely on dynamic differentiable modules
Computation Graph

Automatic computation of gradients: all modules are differentiable!

Tensorflow, theano, etc. build a static computation graph

Torch, pytorch, etc. rely on dynamic differentiable modules

All frameworks enable parallel computation on CPU and GPU
Computation Graph

Simple keras implementation

```python
model = Sequential()
model.add(Dense(H, input_dim=N))  # defines W0
model.add(Activation("tanh"))  # defines W0
model.add(Dense(K))  # defines W1
model.add(Activation("softmax"))
```
Lab 1: here in 15min!