Review
Unit 1
Exercise 1

Theory II

Theory IV

A Deep No

TensorBoard

Building and Training a Deep Learning Network

Deep Learning — Unit 3

Dr. Jon Krohn
jon@untapt.com

Slides available at jonkrohn.com/talks

July 28th, 2018



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory

A Deep Ne

TanaarDaa

- 2 Essential Theory III: Initialization & Mini-Batches
- 3 Essential Theory IV: Unstable Gradients & Overfitting
- 4 A Deep Neural Network
- 5 TensorBoard and the Interpretation of Model Outputs



Heview
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory

A Deep Ne

TensorBoa

- 2 Essential Theory III: Initialization & Mini-Batches
- 3 Essential Theory IV: Unstable Gradients & Overfitting
- 4 A Deep Neural Network
- 5 TensorBoard and the Interpretation of Model Outputs



Heview
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory

A Deep Ne

TensorBoa

- 2 Essential Theory III: Initialization & Mini-Batches
- 3 Essential Theory IV: Unstable Gradients & Overfitting
- 4 A Deep Neural Network
- 5 TensorBoard and the Interpretation of Model Outputs



Ceview
Unit 1
Exercise 1
Unit 2
Exercise 2

-.

A Deep Ne

TensorBoa

- 2 Essential Theory III: Initialization & Mini-Batches
- 3 Essential Theory IV: Unstable Gradients & Overfitting
- 4 A Deep Neural Network
- 5 TensorBoard and the Interpretation of Model Outputs



Ceview
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory

A Deep Ne

TensorBoa

- 2 Essential Theory III: Initialization & Mini-Batches
- 3 Essential Theory IV: Unstable Gradients & Overfitting
- 4 A Deep Neural Network
- 5 TensorBoard and the Interpretation of Model Outputs



Review

Exercise 1
Unit 2
Exercise 2

Theory

A Doon No

T----D--

1 Review Content and Take-Home Exercises

Unit 1

Take-Home Exercise 1

Unit 2

Exercise 2

2 Essential Theory III: Initialization & Mini-Batches

- 3 Essential Theory IV: Unstable Gradients & Overfitting
- 4 A Deep Neural Network
- 5 TensorBoard and the Interpretation of Model Outputs



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

111001 y

A Deep Ne

_ '''

1 Review Content and Take-Home Exercises
Unit 1

Take-Home Exercise 1
Unit 2
Exercise 2

- 2 Essential Theory III: Initialization & Mini-Batches
- 3 Essential Theory IV: Unstable Gradients & Overfitting
- 4 A Deep Neural Network
- 5 TensorBoard and the Interpretation of Model Outputs



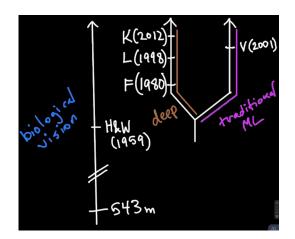
Review
Unit 1
Exercise 1

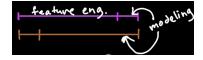
Theory II

Thoony IV

A Deep Ne

TensorBoard







Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

Theory I'

A Deep No

TensorBoard

MNIST Digits & LeNet-5

LeCun, Boutou, Bengio & Haffner (1998)



PROC. OF THE IEEE, NOVEMBER 1998

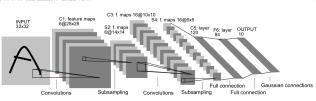


Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.



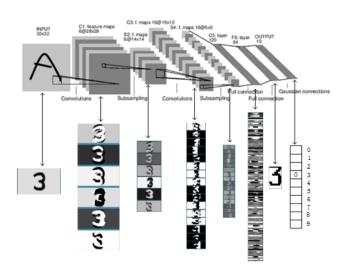
LeNet-5

LeCun, Boutou, Bengio & Haffner (1998)

Unit 1
Exercise 1
Unit 2
Exercise 2
Theory II

A Deep Ne

TensorBoard





Review
Unit 1
Exercise 1
Unit 2

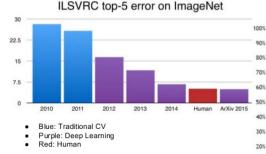
Theory II

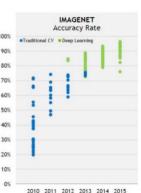
.....

TensorBoard

ImageNet Classification Error

ILSVRC: 1.4m, 1k object classes







Review
Unit 1
Exercise 1
Unit 2

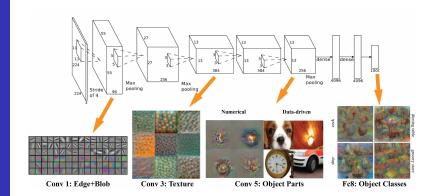
Theory III

A Deep Ne

TensorBoard

AlexNet

Krizhevsky, Sutskever & Hinton (2012)

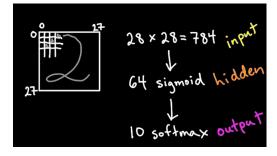




A Shallow Neural Network

Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II



[shallow notebook]



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory I

A Deep Ne

TensorRos

1 Review Content and Take-Home Exercises

Unit 1

Take-Home Exercise 1

Unit 2

Exercise 2

- 2 Essential Theory III: Initialization & Mini-Batches
- 3 Essential Theory IV: Unstable Gradients & Overfitting
- 4 A Deep Neural Network
- 5 TensorBoard and the Interpretation of Model Outputs



TensorFlow Playground

Interactive ANN Visualization

Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

Theory I

A Deep No

TensorBoard

- siamaid nauran
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradient
 - descent

- input laver
- hidden layer
- dense/FC layer
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2

Theory II

Theory I

A Deep Ne

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradient

- input layer
- hidden layer
- dense/FC layer
- softmax layer
- output layer



Unit 1
Exercise 1
Unit 2

Theory II

Theory IV

A Deep No

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
 - ReLU

- weight
- bias
- cost function
- gradien
 - descent

- input laver
- hidden layer
- dense/FC layer
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

/ Doop it

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradien
 - descent

- input laver
- hidden layer
- dense/FC lave
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

,

// Doop ivo

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradien
 - descent

- input laver
- hidden layer
- donco/EC lov
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

A D - - - NI

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradient descent

- input layer
- hidden layer
- dense/FC layer
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

A Doon No

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradien
 descent

- input layer
- hidden layer
- dense/FC layer
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

A Doop No

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradien descent

- input layer
- hidden layer
- dense/FC layer
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradient descent

- input layer
- hidden layer
- dense/FC layer
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradient descent

- input layer
- hidden layer
- dense/FC lave
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

THEOLY IV

...

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradient descent

- input layer
- hidden layer
- dense/FC layer
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradient descent

- input layer
- hidden layer
- dense/FC layer
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

A Doon No

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradient descent

- input layer
- hidden layer
- dense/FC layer
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

THEOLY IV

TopoorPoore

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradient descent

- input layer
- hidden layer
- dense/FC layer
 - softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

Theory IV

A Deep Ne

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradient descent

- input layer
- hidden layer
- dense/FC layer
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

TensorBoard

TensorFlow Playground

- sigmoid neuron
- tanh neuron
- ReLU

- weight
- bias
- cost function
- gradient descent

- input layer
- hidden layer
- dense/FC layer
- softmax layer
- output layer



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Thoory I

A Deep Ne

TopoorPor

Review Content and Take-Home Exercises

Unit 1

Take-Home Exercise 1

Unit 2

Exercise 2

2 Essential Theory III: Initialization & Mini-Batches

- 3 Essential Theory IV: Unstable Gradients & Overfitting
- 4 A Deep Neural Network
- 5 TensorBoard and the Interpretation of Model Outputs



Essential Theory I Neural Units

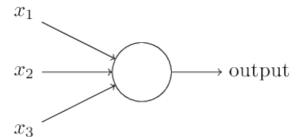
Review
Unit 1
Exercise 1
Unit 2

Theory II

Theory I

A Deep No

TensorBoard





Review
Unit 1
Exercise 1
Unit 2

Theory II

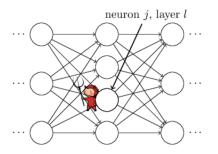
Theory is

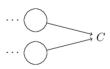
v peeb Me

TensorBoard

Essential Theory II

Cost Functions, Gradient Descent, and Backpropagation







An Intermediate Neural Network

Review
Unit 1

Exercise 1
Unit 2
Exercise 2

Theory II

Theory I

A Deep No

TensorBoard

[intermediate notebook]



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory I

A Deep Ne

TensorBoar

Data Sets for Deep Learning





| Bluebell | | |
|-------------|---|------|
| O Committee | 1 | |
| Tigerlify | | |
| Tulip | | W We |
| Sowslip | | 1 |

| Dataset | Classes | Train Samples |
|------------------------|---------|---------------|
| AG's News | 4 | 120,000 |
| Sogou News | 5 | 450,000 |
| DBPedia | 14 | 560,000 |
| Yelp Review Polarity | 2 | 560,000 |
| Yelp Review Full | 5 | 650,000 |
| Yahoo! Answers | 10 | 1,400,000 |
| Amazon Review Full | 5 | 3,000,000 |
| Amazon Review Polarity | 2 | 3,600,000 |



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory I

A Deep Ne

TensorBoard

Data Sets for Deep Learning





| Bluchell | |
|----------|-----|
| Crosus | |
| Treerliv | |
| Tulip (| *** |
| disease) | |

| Dataset | Classes | Train Samples |
|------------------------|---------|---------------|
| AG's News | 4 | 120,000 |
| Sogou News | 5 | 450,000 |
| DBPedia | 14 | 560,000 |
| Yelp Review Polarity | 2 | 560,000 |
| Yelp Review Full | 5 | 650,000 |
| Yahoo! Answers | 10 | 1,400,000 |
| Amazon Review Full | 5 | 3,000,000 |
| Amazon Review Polarity | 2 | 3,600,000 |



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory I

A Deep Ne

TensorBoard

Data Sets for Deep Learning





| Bluchell | |
|----------|-----|
| Crosus | |
| Treerliv | |
| Tulip (| *** |
| disease) | |

| Dataset | Classes | Train Samples |
|------------------------|---------|---------------|
| AG's News | 4 | 120,000 |
| Sogou News | 5 | 450,000 |
| DBPedia | 14 | 560,000 |
| Yelp Review Polarity | 2 | 560,000 |
| Yelp Review Full | 5 | 650,000 |
| Yahoo! Answers | 10 | 1,400,000 |
| Amazon Review Full | 5 | 3,000,000 |
| Amazon Review Polarity | 2 | 3,600,000 |



Review
Unit 1
Exercise 1

Unit 2 Exercise 2

Theory II

111001 y 1 v

A Deep Ne

TensorBoard





Jon Krohn, Cajoler of Datums

Home Resources

Posts Publications

Talks

Academia Applications Quotations

Open Data Sources

To train a powerful model, the larger the data set, the better -- if it's well-organised and open, that's ideal. The following repositories are standouts that meet all these criteria:

- Data.gov (home of >150k US government-related datasets),
- Govcode, a collection of government open source projects,
- the Open Data Stack Exchange, and
- · this curated list of 'awesome' public datasets
- this well-annotated list of data sets for natural language processing
- for biomedical and health data specifically, check out:
 - this University of Minnesota resource
 - this Medical Data for Machine Learning GitHub repo

For machine learning models that require a lot of labelled data, check out:

- UC Irvine's repository
- · Yahoo's massive 13TB data set comprised of 100 billion user interactions with news items
- Google's image and video data sets
- Luke de Oliveira's Greatest Public Datasets for Al blog post
- CrowdFlower's Data for Everyone

Finally, here are extensive pages on importing data from the Web into R, provided by CRAN and MRAN.



Outline

Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory

A Deep Ne

TanaarDaa

1 Review Content and Take-Home Exercises

Unit 1

Take-Home Exercise 1

Unit 2

Exercise 2

2 Essential Theory III: Initialization & Mini-Batches

3 Essential Theory IV: Unstable Gradients & Overfitting

4 A Deep Neural Network

5 TensorBoard and the Interpretation of Model Outputs



Your Deep Learning Project I

Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

A Deep Ne







Theory II

A Door No

TensorBoard

Your Deep Learning Project I

Perspectives to approach ideating from:

- Identify a data set ⇒ use it to solve a problem
- Identify a problem that you'd like to solve ⇒ find an appropriate data set



Theory II

Theory IV

A Deep No

TensorBoard

Your Deep Learning Project I

Perspectives to approach ideating from:

- Identify a data set ⇒ use it to solve a problem
- Identify a problem that you'd like to solve ⇒ find an appropriate data set



Theory II

A Deep No

TensorBoard

- 1 a machine-vision architecture to classify images, e.g.:
 - [Fashion MNIST]
 - one of the dozens of data sets with the keyword image in the title from [CrowdFlower]
 - one of the Computer Vision data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 2 an NLP architecture to classify text, e.g.:
 - the Yelp or Amazon sentiment [data sets] detailed in Section 4 of [Xiang Zhang et al.'s paper]
 - the Yahoo! Answers categories data set also detailed in Xiang Zhang et al.'s paper
 - one of the dozens of data sets with the keywords sentiment or text in the title from [CrowdFlower]
 - one of the Natural Language data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 3 a regression mode



Theory II

A Deep Ne

TensorBoar

- 1 a machine-vision architecture to classify images, e.g.:
 - [Fashion MNIST]
 - one of the dozens of data sets with the keyword image in the title from [CrowdFlower]
 - one of the Computer Vision data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 2 an NLP architecture to classify text, e.g.:
 - the Yelp or Amazon sentiment [data sets] detailed in Section 4 of [Xiang Zhang et al.'s paper]
 - the Yahoo! Answers categories data set also detailed in Xiang Zhang et al.'s paper
 - one of the dozens of data sets with the keywords sentiment or text in the title from [CrowdFlower]
 - one of the Natural Language data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 3 a regression mode



Theory II
Theory IV

A Deep Ne

TensorBoard

- 1 a machine-vision architecture to classify images, e.g.:
 - [Fashion MNIST]
 - one of the dozens of data sets with the keyword image in the title from [CrowdFlower]
 - one of the Computer Vision data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- ② an NLP architecture to classify text, e.g.:
 - the Yelp or Amazon sentiment [data sets] detailed in Section 4 of [Xiang Zhang et al.'s paper]
 - the Yahoo! Answers categories data set also detailed in Xiang Zhang et al.'s paper
 - one of the dozens of data sets with the keywords sentiment or text in the title from [CrowdFlower]
 - one of the Natural Language data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 3 a regression mode



Theory II

A Doon No

TensorBoard

- 1 a machine-vision architecture to classify images, e.g.:
 - [Fashion MNIST]
 - one of the dozens of data sets with the keyword image in the title from [CrowdFlower]
 - one of the Computer Vision data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 2 an NLP architecture to classify text, e.g.:
 - the Yelp or Amazon sentiment [data sets] detailed in Section 4 of [Xiang Zhang et al.'s paper]
 - the Yahoo! Answers categories data set also detailed in Xiang Zhang et al.'s paper
 - one of the dozens of data sets with the keywords sentiment or text in the title from [CrowdFlower]
 - one of the Natural Language data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 3 a regression mode



Theory II

A Doop No

TensorBoard

- 1 a machine-vision architecture to classify images, e.g.:
 - [Fashion MNIST]
 - one of the dozens of data sets with the keyword image in the title from [CrowdFlower]
 - one of the Computer Vision data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 2 an NLP architecture to classify text, e.g.:
 - the Yelp or Amazon sentiment [data sets] detailed in Section 4 of [Xiang Zhang et al.'s paper]
 - the Yahoo! Answers categories data set also detailed in Xiang Zhang et al.'s paper
 - one of the dozens of data sets with the keywords sentiment or text in the title from [CrowdFlower]
 - one of the Natural Language data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- a regression mode



Theory II

A Deep Ne

TensorBoard

- 1 a machine-vision architecture to classify images, e.g.:
 - [Fashion MNIST]
 - one of the dozens of data sets with the keyword image in the title from [CrowdFlower]
 - one of the Computer Vision data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 2 an NLP architecture to classify text, e.g.:
 - the Yelp or Amazon sentiment [data sets] detailed in Section 4 of [Xiang Zhang et al.'s paper]
 - the Yahoo! Answers categories data set also detailed in Xiang Zhang et al.'s paper
 - one of the dozens of data sets with the keywords sentiment or text in the title from [CrowdFlower]
 - one of the Natural Language data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 3 a regression mode



Theory II

A Deep Ne

TensorBoard

- 1 a machine-vision architecture to classify images, e.g.:
 - [Fashion MNIST]
 - one of the dozens of data sets with the keyword image in the title from [CrowdFlower]
 - one of the Computer Vision data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 2 an NLP architecture to classify text, e.g.:
 - the Yelp or Amazon sentiment [data sets] detailed in Section 4 of [Xiang Zhang et al.'s paper]
 - the Yahoo! Answers categories data set also detailed in Xiang Zhang et al.'s paper
 - one of the dozens of data sets with the keywords sentiment or text in the title from [CrowdFlower]
 - one of the Natural Language data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 3 a regression mode



Theory II

A Deep Ne

TensorBoard

- 1 a machine-vision architecture to classify images, e.g.:
 - [Fashion MNIST]
 - one of the dozens of data sets with the keyword image in the title from [CrowdFlower]
 - one of the Computer Vision data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 2 an NLP architecture to classify text, e.g.:
 - the Yelp or Amazon sentiment [data sets] detailed in Section 4 of [Xiang Zhang et al.'s paper]
 - the Yahoo! Answers categories data set also detailed in Xiang Zhang et al.'s paper
 - one of the dozens of data sets with the keywords sentiment or text in the title from [CrowdFlower]
 - one of the Natural Language data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- a regression mode



Theory II

A Deep Ne

TensorBoard

- 1 a machine-vision architecture to classify images, e.g.:
 - [Fashion MNIST]
 - one of the dozens of data sets with the keyword image in the title from [CrowdFlower]
 - one of the Computer Vision data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 2 an NLP architecture to classify text, e.g.:
 - the Yelp or Amazon sentiment [data sets] detailed in Section 4 of [Xiang Zhang et al.'s paper]
 - the Yahoo! Answers categories data set also detailed in Xiang Zhang et al.'s paper
 - one of the dozens of data sets with the keywords sentiment or text in the title from [CrowdFlower]
 - one of the Natural Language data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- a regression mode



Theory II

A Deep Ne

TensorBoard

- 1 a machine-vision architecture to classify images, e.g.:
 - [Fashion MNIST]
 - one of the dozens of data sets with the keyword image in the title from [CrowdFlower]
 - one of the Computer Vision data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 2 an NLP architecture to classify text, e.g.:
 - the Yelp or Amazon sentiment [data sets] detailed in Section 4 of [Xiang Zhang et al.'s paper]
 - the Yahoo! Answers categories data set also detailed in Xiang Zhang et al.'s paper
 - one of the dozens of data sets with the keywords sentiment or text in the title from [CrowdFlower]
 - one of the Natural Language data sets (other than the MNIST data set) in Luke de Oliveira's [blog post]
- 3 a regression model



Outline

Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory III

A Deep Ne

TensorRos

1 Review Content and Take-Home Exercises
Unit 1
Take-Home Exercise 1
Unit 2
Exercise 2

- 2 Essential Theory III: Initialization & Mini-Batches
- 3 Essential Theory IV: Unstable Gradients & Overfitting
- 4 A Deep Neural Network
- 5 TensorBoard and the Interpretation of Model Outputs



Weight Initialization

Review

Exercise 1 Unit 2 Exercise 2

Theory III

Thooly IV

Y Deeb M

ToncorBoar

- uniform
- normal
- Xavier Glorot

[Jupyter demo]



Weight Initialization

Review

Exercise 1 Unit 2 Exercise 2

Theory III

THEOLYTV

... Воор ...

- -

- uniform
- normal
- Xavier Glorot

[Jupyter demo]





Weight Initialization

Review

Exercise 1
Unit 2
Exercise 2

Theory III

Theory IV

A Deep Ne

TensorBoard

- uniform
- normal
- Xavier Glorot

[Jupyter demo]



Unit 1 Exercise 1 Unit 2 Exercise 2

Theory III

A Door No

- · learning rate
- batch size
- second-order gradient learning
 - momentum
 - Adam



Unit 1 Exercise 1 Unit 2 Exercise 2

Theory III

A D. . . . NI.

- · learning rate
- batch size
- second-order gradient learning
 - momentum
 - Adam



Unit 1
Exercise 1
Unit 2
Exercise 2

Theory III

A Deep No

ToncorBook

- · learning rate
- batch size
- second-order gradient learning
 - momentum
 - Adam



Unit 1 Exercise 1 Unit 2 Exercise 2

Theory III

A Doop N

- · learning rate
- batch size
- · second-order gradient learning
 - momentum
 - Adam



Unit 1 Exercise 1 Unit 2 Exercise 2

Theory III

A Deep Ne

- · learning rate
- batch size
- second-order gradient learning
 - momentum
 - Adam



Outline

Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory

Theory IV

71 Doop 140

1 Review Content and Take-Home Exercises
Unit 1
Take-Home Exercise 1
Unit 2
Exercise 2

- 2 Essential Theory III: Initialization & Mini-Batches
- 3 Essential Theory IV: Unstable Gradients & Overfitting
- 4 A Deep Neural Network
- 5 TensorBoard and the Interpretation of Model Outputs



Review

Unit 1
Exercise 1
Unit 2
Exercise 2

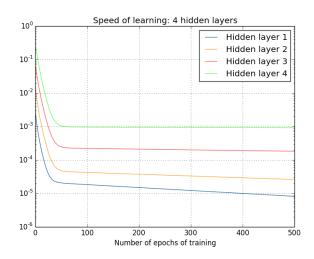
Theory II

Theory IV

A Deep No

TensorBoard

Unstable Gradients





Unstable Gradients

Review

Exercise ' Unit 2

Theory II

Theory IV

A Deep No

- vanishing
- exploding



Unstable Gradients

Review

Exercise 1 Unit 2

Theory II

Theory IV

A Deep No

- vanishing
- exploding



Review
Unit 1

Exercise 1 Unit 2 Exercise 2

Theory II

Theory IV

y peeb ive

TensorBoard

Avoiding Overfitting

(Or, Model Generalization)

- L1/L2 regularization
- dropout
- artificial data set expansion



Review

Exercise 1
Unit 2

Theory II

Theory IV

A Deep Ne

TensorBoard

Avoiding Overfitting

(Or, Model Generalization)

- L1/L2 regularization
- dropout
- artificial data set expansion



Review

Exercise 1 Unit 2 Exercise 2

Theory II

Theory IV

TensorBoar

Avoiding Overfitting

(Or, Model Generalization)

- L1/L2 regularization
- dropout
- artificial data set expansion



Avoiding Overfitting

(Or, Model Generalization)

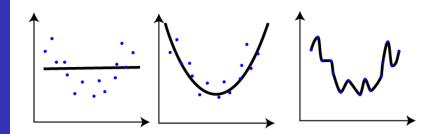
Heview
Unit 1

Exercise 1 Unit 2 Exercise 2

Theory II

Theory IV

A Deep Ne





Last, But Not Least

Unit 1 Exercise 1 Unit 2

Theory II

Theory IV

A Deep Ne

- more layers
 - max-pooling
 - flatten
- batch normalization avoids covariate shift; advantages:
 - initialization parameters
 - avoid neuron saturation
 - g regularizing effect



Last, But Not Least

Unit 1 Exercise 1 Unit 2

Theory IV

A Deep Ne

... Воор ...

- more layers
 - max-pooling
 - flatten
- batch normalization avoids covariate shift; advantages:
 - initialization parameters
 - avoid neuron saturation
 - g regularizing effect



Unit 1 Exercise 1 Unit 2 Exercise 2

Theory IV

A Deep Ne

- more layers
 - max-pooling
 - flatten
- batch normalization avoids covariate shift; advantages:
 - initialization parameters
 - avoid neuron saturation
 - regularizing effect



Unit 1
Exercise 1
Unit 2

Theory IV

A Deep No

- more layers
 - max-pooling
 - flatten
- batch normalization avoids covariate shift; advantages:
 - initialization parameters
 - avoid neuron saturation
 - 3 regularizing effect



Unit 1 Exercise 1 Unit 2 Exercise 2

Theory IV

A Deen Ne

...---

- more layers
 - max-pooling
 - flatten
- batch normalization avoids covariate shift; advantages:
 - 1 initialization parameters
 - avoid neuron saturation
 - 3 regularizing effect



Unit 1
Exercise 1
Unit 2
Exercise 2

Theory IV

A Doop No

A Deep N

- more layers
 - max-pooling
 - flatten
- batch normalization avoids covariate shift; advantages:
 - 1 initialization parameters
 - 2 avoid neuron saturation
 - 3 regularizing effect



Theory IV

- more layers
 - max-pooling
 - flatten
- batch normalization avoids covariate shift; advantages:
 - initialization parameters
 - avoid neuron saturation
 - 3 regularizing effect



Outline

Review
Unit 1
Exercise 1
Unit 2
Exercise 2

A Deep Net

. . Воор . . о

rensorboar

1 Review Content and Take-Home Exercises
Unit 1
Take-Home Exercise 1
Unit 2
Exercise 2

- 2 Essential Theory III: Initialization & Mini-Batches
- 3 Essential Theory IV: Unstable Gradients & Overfitting
- 4 A Deep Neural Network
- 5 TensorBoard and the Interpretation of Model Outputs



Unit 3 — Building a Deep Net

Reviev

Exercise 1

Theory II

Theory I\

A Deep Net

TensorBoard

Let's make [intermediate net] deep!



Outline

Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory I

A Deep Ne

TensorBoard

1 Review Content and Take-Home Exercises
Unit 1
Take-Home Exercise 1
Unit 2
Exercise 2

- 2 Essential Theory III: Initialization & Mini-Batches
- 3 Essential Theory IV: Unstable Gradients & Overfitting
- 4 A Deep Neural Network
- 5 TensorBoard and the Interpretation of Model Outputs



Unit 1 Exercise 1 Unit 2 Exercise 2

Theory II

A Doop No

A Deep Ne

- add callback as in [Deep Net in Keras Jupyter notebook]
- ② use Terminal to navigate to your logs directory
- 3 run tensorboard --logdir=. --port 6006
- 4 navigate to http://localhost:6006/ in a web browser



Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

A Deen No

- add callback as in [Deep Net in Keras Jupyter notebook]
- 2 use Terminal to navigate to your logs directory
- 3 run tensorboard --logdir=. --port 6006
- 4 navigate to http://localhost:6006/ in a web browser



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

A Doop No

- add callback as in [Deep Net in Keras Jupyter notebook]
- 2 use Terminal to navigate to your logs directory
- 3 run tensorboard --logdir=. --port 6006
- 4 navigate to http://localhost:6006/ in a web browser



Review
Unit 1
Exercise 1
Unit 2
Exercise 2

Theory II

A Doop No

- add callback as in [Deep Net in Keras Jupyter notebook]
- 2 use Terminal to navigate to your logs directory
- 3 run tensorboard --logdir=. --port 6006
- 4 navigate to http://localhost:6006/ in a web browser



Unit 3 — Building a Deep Net

Review Unit 1 Exercise 1 Unit 2

Theory III

Theory II

A Deen Ne

TensorBoard

TensorBoard

The Interpretation of Model Outputs

