

# How Deep Learning Works

## Deep Learning — Unit 2

Dr. Jon Krohn

`jon@untapt.com`

Slides available at `jonkrohn.com/talks`

October 20th, 2018

# Outline

- 1 Essential Theory I: Neural Units
- 2 Essential Theory II: Cost Functions, Gradient Descent & Backpropagation
- 3 An Intermediate Neural Network
- 4 Data Sets for Deep Learning
- 5 Your Deep Learning Project I: Ideating

# Outline

- 1 Essential Theory I: Neural Units
- 2 Essential Theory II: Cost Functions, Gradient Descent & Backpropagation
- 3 An Intermediate Neural Network
- 4 Data Sets for Deep Learning
- 5 Your Deep Learning Project I: Ideating

# Outline

- 1 Essential Theory I: Neural Units
- 2 Essential Theory II: Cost Functions, Gradient Descent & Backpropagation
- 3 An Intermediate Neural Network
- 4 Data Sets for Deep Learning
- 5 Your Deep Learning Project I: Ideating

# Outline

- 1 Essential Theory I: Neural Units
- 2 Essential Theory II: Cost Functions, Gradient Descent & Backpropagation
- 3 An Intermediate Neural Network
- 4 Data Sets for Deep Learning
- 5 Your Deep Learning Project I: Ideating

# Outline

- 1 Essential Theory I: Neural Units
- 2 Essential Theory II: Cost Functions, Gradient Descent & Backpropagation
- 3 An Intermediate Neural Network
- 4 Data Sets for Deep Learning
- 5 Your Deep Learning Project I: Ideating

# Outline

- 1 Essential Theory I: Neural Units
- 2 Essential Theory II: Cost Functions, Gradient Descent & Backpropagation
- 3 An Intermediate Neural Network
- 4 Data Sets for Deep Learning
- 5 Your Deep Learning Project I: Ideating

Neuron  
Theory

Network  
Theory

Intermediate  
Net

Data for Deep  
Learning

Your Project

[whiteboard equations + Arsenal]



Neuron  
Theory

Network  
Theory

Intermediate  
Net

Data for Deep  
Learning

Your Project

[playground]

# Outline

- 1 Essential Theory I: Neural Units
- 2 Essential Theory II: Cost Functions, Gradient Descent & Backpropagation**
- 3 An Intermediate Neural Network
- 4 Data Sets for Deep Learning
- 5 Your Deep Learning Project I: Ideating

[whiteboard equations + Arsenal]

Neuron  
Theory

**Network  
Theory**

Intermediate  
Net

Data for Deep  
Learning

Your Project

[playground]

# Take-Home Exercise I

## Artificial Neural Network Language

- solve another TF Playground problem, e.g., a regression
- verbalize weights, biases & network architecture
- verbalize Arsenal items

# Take-Home Exercise I

## Artificial Neural Network Language

- solve another TF Playground problem, e.g., a regression
- verbalize weights, biases & network architecture
- verbalize Arsenal items

# Take-Home Exercise I

## Artificial Neural Network Language

- solve another TF Playground problem, e.g., a regression
- verbalize weights, biases & network architecture
- verbalize Arsenal items

# Outline

- 1 Essential Theory I: Neural Units
- 2 Essential Theory II: Cost Functions, Gradient Descent & Backpropagation
- 3 An Intermediate Neural Network**
- 4 Data Sets for Deep Learning
- 5 Your Deep Learning Project I: Ideating



Neuron  
Theory

Network  
Theory

**Intermediate  
Net**

Data for Deep  
Learning

Your Project

[how might you...?]

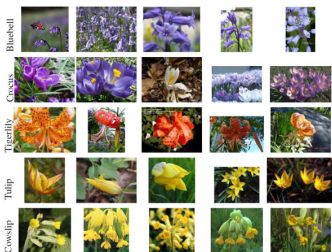
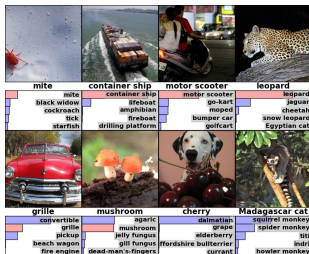
# Outline

- 1 Essential Theory I: Neural Units
- 2 Essential Theory II: Cost Functions, Gradient Descent & Backpropagation
- 3 An Intermediate Neural Network
- 4 Data Sets for Deep Learning**
- 5 Your Deep Learning Project I: Ideating

Neuron  
TheoryNetwork  
TheoryIntermediate  
NetData for Deep  
Learning

Your Project

0	4	1	9	2	1	3	1	4	3
5	3	6	1	7	2	8	6	9	4
0	9	1	1	2	4	3	2	7	3
8	6	9	0	5	6	0	7	6	1
8	7	9	3	9	8	5	9	3	3
0	7	4	9	8	0	9	4	7	4
4	6	0	4	5	6	1	0	0	1
7	1	6	3	0	2	7	7	7	9
0	2	6	7	8	3	9	0	4	6
7	4	6	8	0	7	8	3	7	5

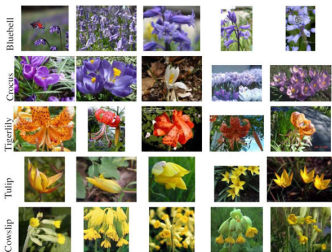
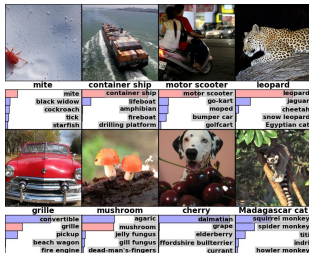


Dataset	Classes	Train Samples
AG's News	4	120,000
Sogou News	5	450,000
DBPedia	14	560,000
Yelp Review Polarity	2	560,000
Yelp Review Full	5	650,000
Yahoo! Answers	10	1,400,000
Amazon Review Full	5	3,000,000
Amazon Review Polarity	2	3,600,000

Neuron  
TheoryNetwork  
TheoryIntermediate  
NetData for Deep  
Learning

Your Project

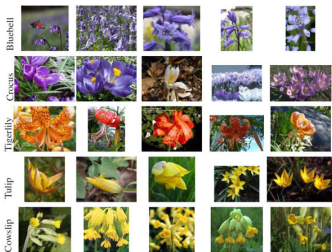
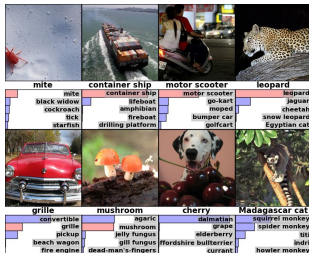
0	4	1	9	2	1	3	1	4	3
5	3	6	1	7	2	8	6	9	4
0	9	1	1	2	4	3	2	7	3
8	6	9	0	5	6	0	7	6	1
8	7	9	3	9	8	5	9	3	3
0	7	4	9	8	0	9	4	7	4
4	6	0	4	5	6	1	0	0	1
7	1	6	3	0	2	7	7	7	9
0	2	6	7	8	3	9	0	4	6
7	4	6	8	0	7	8	3	7	5



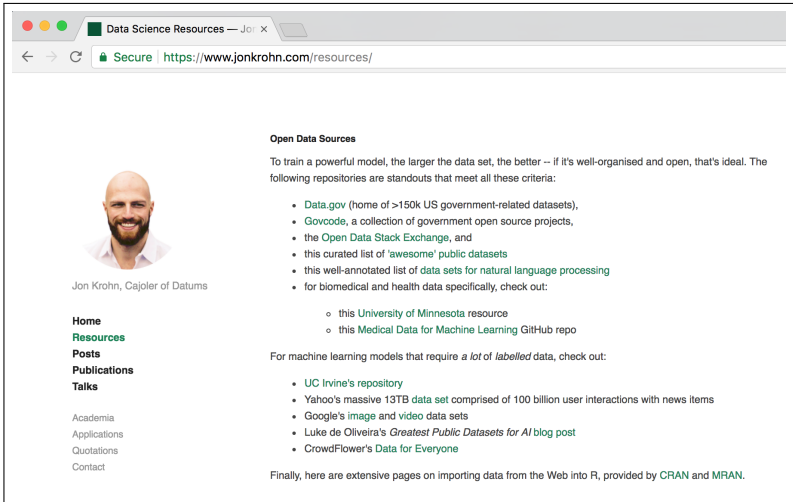
Dataset	Classes	Train Samples
AG's News	4	120,000
Sogou News	5	450,000
DBPedia	14	560,000
Yelp Review Polarity	2	560,000
Yelp Review Full	5	650,000
Yahoo! Answers	10	1,400,000
Amazon Review Full	5	3,000,000
Amazon Review Polarity	2	3,600,000

Neuron  
TheoryNetwork  
TheoryIntermediate  
NetData for Deep  
Learning

Your Project




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



Data Science Resources — Jon Krohn

Secure | <https://www.jonkrohn.com/resources/>



Jon Krohn, Cajoler of Datums

**Home**

**Resources**

**Posts**

**Publications**

**Talks**

Academia

Applications

Quotations

Contact

**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 AI](#) 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](#).

[Data Is Plural]

# Outline

- 1 Essential Theory I: Neural Units
- 2 Essential Theory II: Cost Functions, Gradient Descent & Backpropagation
- 3 An Intermediate Neural Network
- 4 Data Sets for Deep Learning
- 5 Your Deep Learning Project I: Ideating

# Your Deep Learning Project I

## Ideating

Neuron  
Theory

Network  
Theory

Intermediate  
Net

Data for Deep  
Learning

Your Project





# Your Deep Learning Project I

## Ideating

Neuron  
Theory

Network  
Theory

Intermediate  
Net

Data for Deep  
Learning

Your Project

Perspectives to approach ideating from:

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

# Your Deep Learning Project I

## Ideating

Neuron  
Theory

Network  
Theory

Intermediate  
Net

Data for Deep  
Learning

Your Project

Perspectives to approach ideating from:

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

## Recommended starting points:

- 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

## Recommended starting points:

- 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

## Recommended starting points:

- 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

## Recommended starting points:

- 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

## Recommended starting points:

- 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

## Recommended starting points:

- 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



## Recommended starting points:

- 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

## Recommended starting points:

- 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

## Recommended starting points:

- 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

## Recommended starting points:

- 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