Intro to Deep Learning with Theano and OpenDeep

Markus Beissinger
@mbeissinger
markus@vitruvianscience.com
code available: https://github.com/mbeissinger/odsc

ODSC West Workshop 2015

Outline

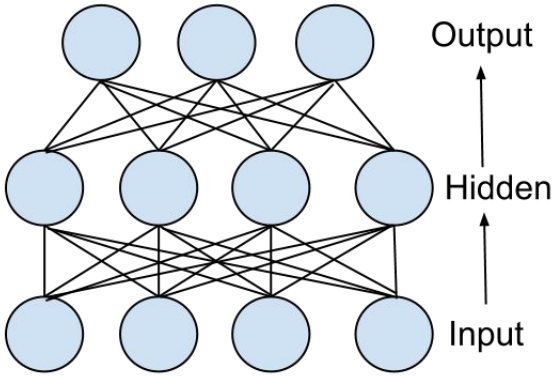
- What is deep learning
- Linear algebra refresher
- Object classification problem
- MLP - most basic neural net
- LeNet - simple convolutional neural net

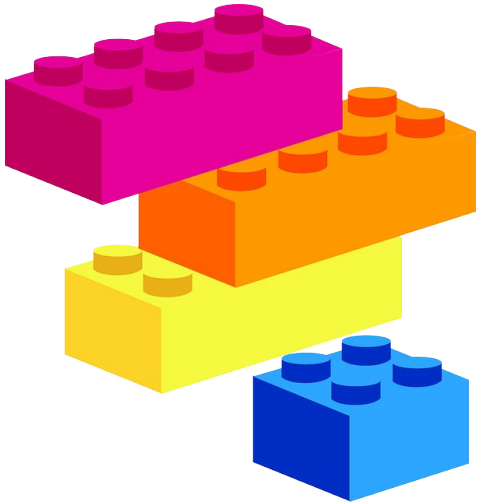
What is deep learning?

- Loosely based on brain - hierarchical feature representations
- Learns useful* representations

*Explains variations in input data with increasingly complex features (edges/curves → nose/eyes/mouth → faces)

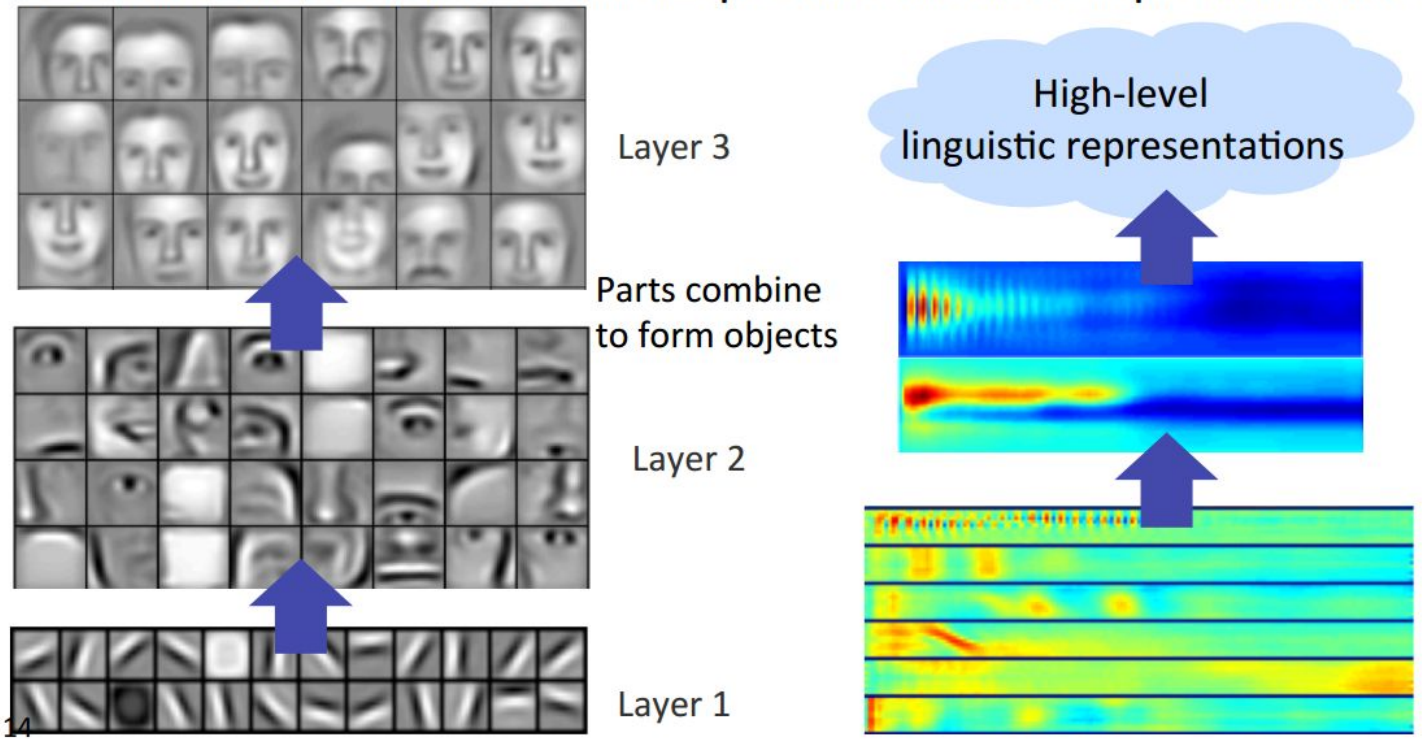
- Automates feature engineering!
- How? Compositions of nonlinear transformations! (stay tuned)





What is deep learning?

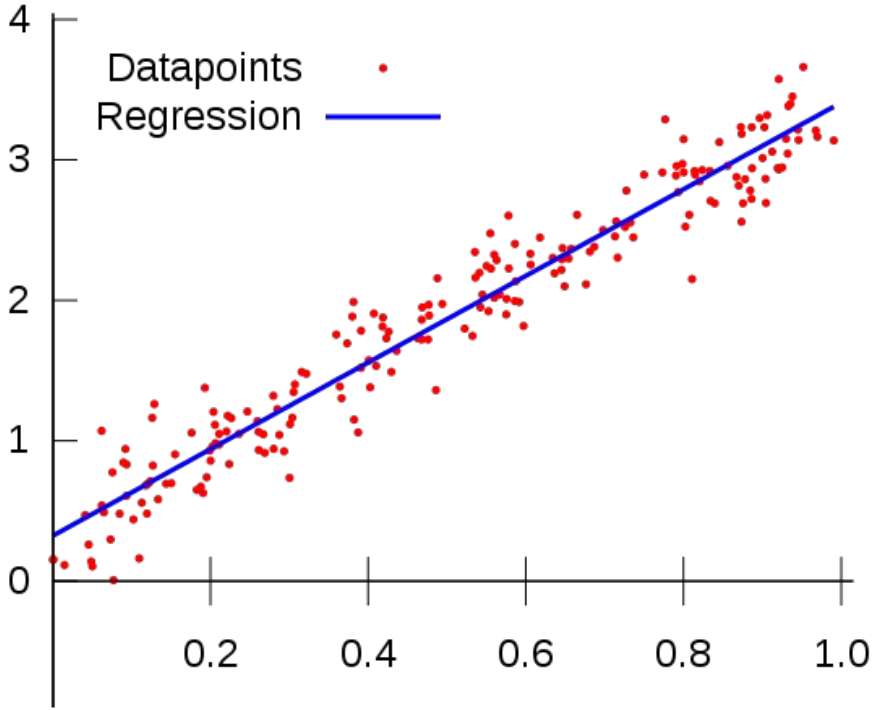
Successive model layers learn deeper intermediate representations



Prior: underlying factors & concepts compactly expressed w/ multiple levels of abstraction

http://www.iro.umontreal.ca/~bengioy/talks/mlss-austin.pdf

Linear algebra refresher: linear regression



$$\hat{y} = mx + b$$

output

slope (parameter)

input

intercept (bias)

$$\hat{y} = W_1x_1 + W_2x_2 + W_3x_3 + b$$

$$\hat{y} = [W_1\ W_2\ W_3] \begin{vmatrix} x_1 \\ x_2 \\ x_3 \end{vmatrix} + b$$

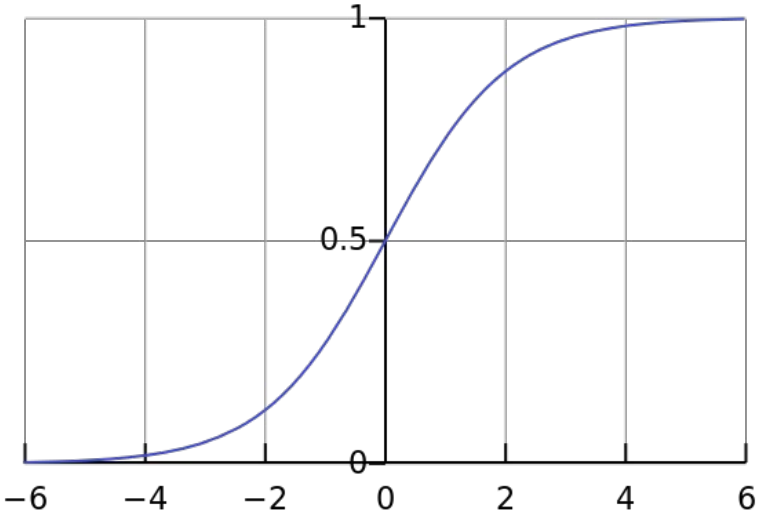
Output = matrix multiplication (linear transform of input)!

Linear algebra refresher: logistic regression

$$\hat{y} = \sigma([W_1\ W_2\ W_3] \begin{vmatrix} x_1 \\ x_2 \\ x_3 \end{vmatrix} + b)$$

$$\hat{y} = \sigma(Wx+b)$$

$$\hat{y} = \frac{1}{1 + e^{-(Wx+b)}}$$



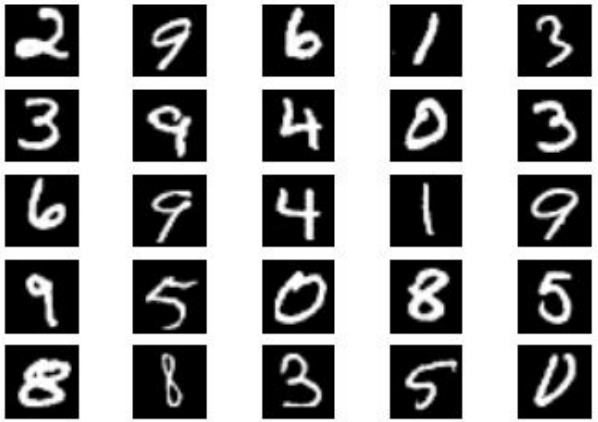
σ = sigmoid logistic function

Output is squashed to [0,1]
now it is a probability!

Object classification

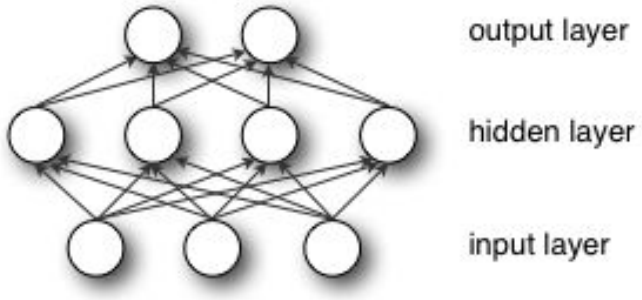
Given an input image, assign a correct class label *for the whole image*.

MNIST handwritten digits



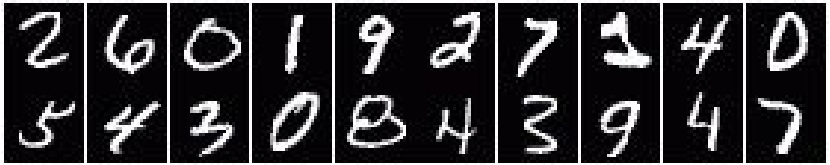
ImageNet food subset

Neural Network: MLP



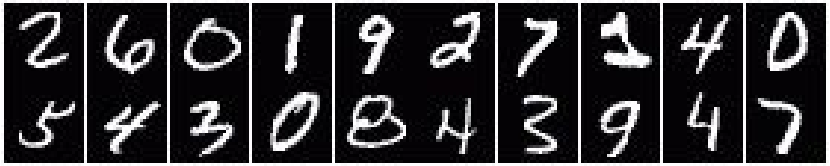
Performs multiple logistic regressions to compute probabilities for each output class!

Neural Network: MLP digit classification



Classify MNIST handwritten digits into one of 10 classes.

Neural Network: MLP architecture



Architecture:

- Input: x = (1, 28*28) vectors
- Hidden layer w/500 hidden units: $H = \sigma(xW_x + b_h)$ W_x= (28*28, 500) weights matrix, b_h=(1, 500) bias vector
- Output layer w/10 output classes: $Y = \text{softmax}(HW_H + b_y)$ W_H=(500, 10) weights matrix, b_y=(1, 10) bias vector

Neural Network: MLP architecture

[code]

Neural Network: MLP training

Cost function:

Negative mean log likelihood of class probabilities P ($Y=y_i|X$).

Because we use softmax, output Y from the MLP is the vector of class probabilities!

cost = -mean(log(Y[index_of_correct_label]))

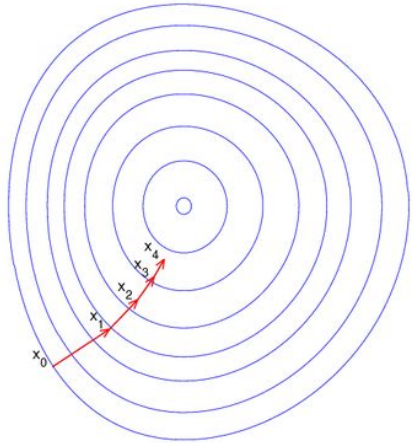
[0.4, .1, .5] -> 2

Neural Network: MLP training

Training:

Stochastic gradient descent (with momentum).

Iterates over training data and makes small changes to weights and biases to maximize log likelihood.



Neural Network: MLP training

[code]

Neural Network: MLP training

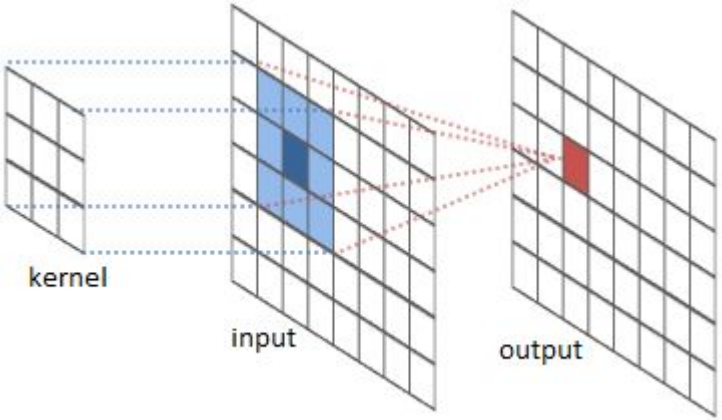
Regularization:

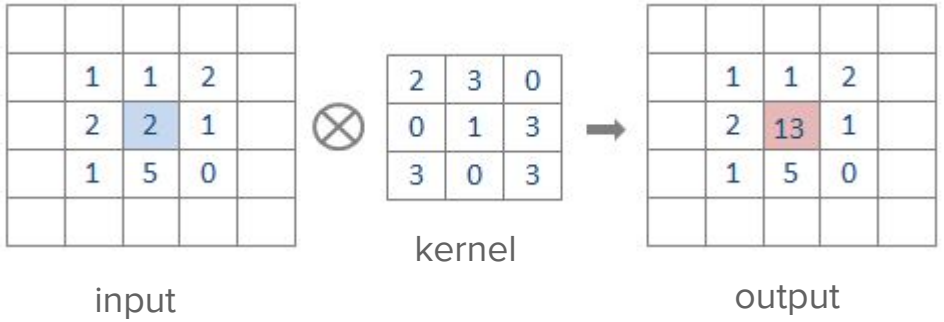
Dropout (masking noise on input and hidden units during training).

Necessary to prevent overfitting and bad generalization to real-world data.

Object Classification - convolutions

Convolutional nets (CNN) use the convolution operator to compute *local* features that take in surrounding context.





images from RiverTrail docs: http://intellabs.github.io/RiverTrail/tutorial/

Convolutional nets

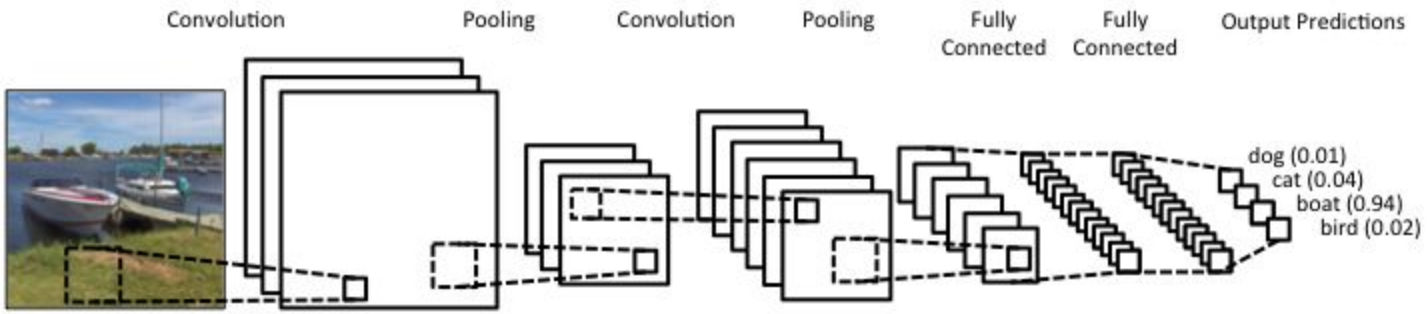
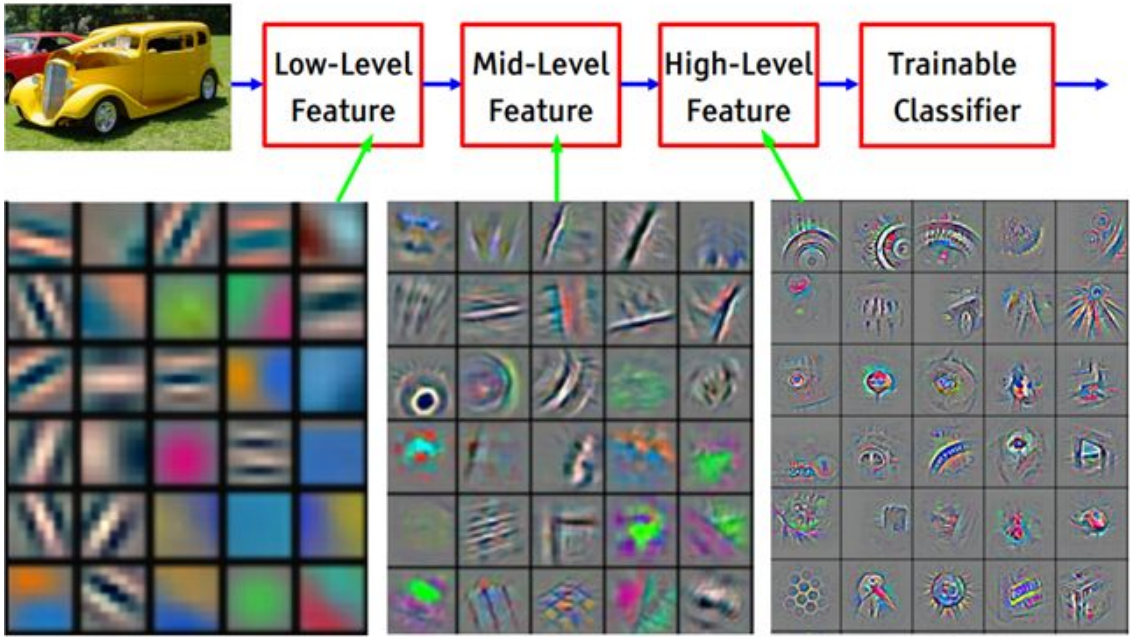


image from http://www.clarifai.com/technology

Convolutional nets

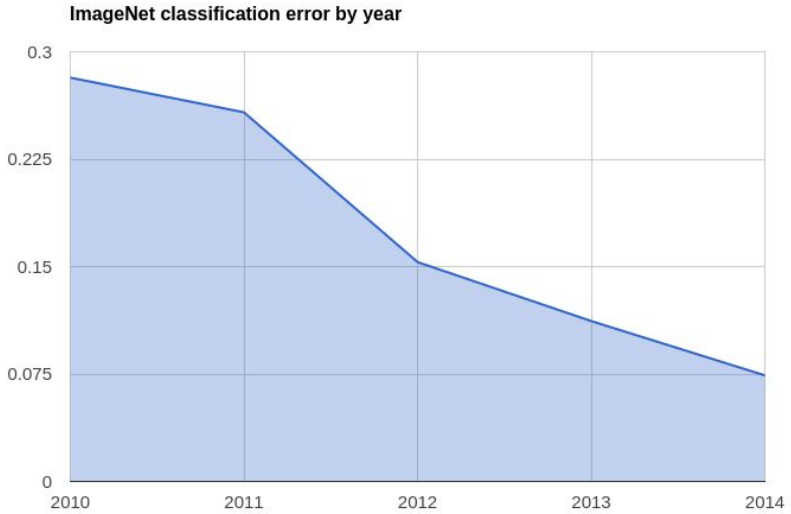


http://www.rsipvision.com/wp-content/uploads/2015/04/Slide8.png

Object classification

Very good results (since 2012) with deep convolutional nets

~7.4% error



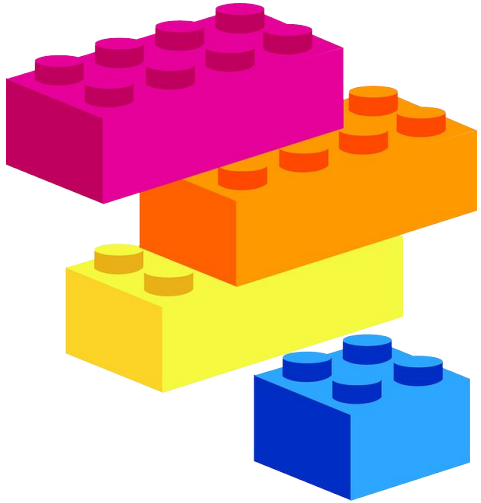
Convolutional nets

[code]

In summary...

What is deep learning?

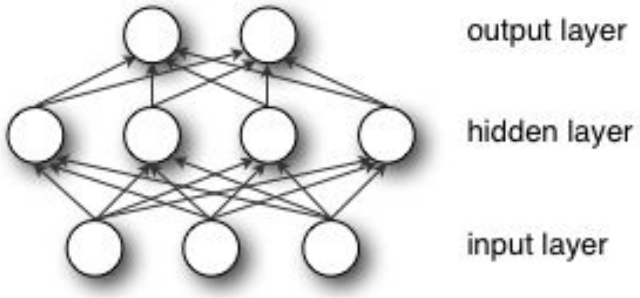
- Model that learns hierarchical representation of input
- How? Uses compositions of nonlinear functions (layers) that learn different levels of abstraction.
- Learns useful features! (no more hand-coded feature engineering)*



*needs lots of data!

What is deep learning?

- Generalizes logistic regressions.
- Uses simplest building blocks (fully-connected layers):
  - Matrix multiplication
  - Nonlinear activation function (ends with softmax)



What is deep learning?

- Deep learning forms a hierarchical representation of input (feature learning!)
- Basic building blocks (layers) are some matrix multiplication + nonlinear activation
- Complex networks made with simple building blocks -> all trained with one algorithm using gradients + chain rule.

Thank you!

Questions?

@mbeissinger
markus@vitruvianscience.com
code available: https://github.com/mbeissinger/odsc