MACHINE LEARNING IN HIGH ENERGY PHYSICS

PRACTICAL CLASS #3

Alex Rogozhnikov, 2015

LINEAR MODELS

Work fine in applications with many features

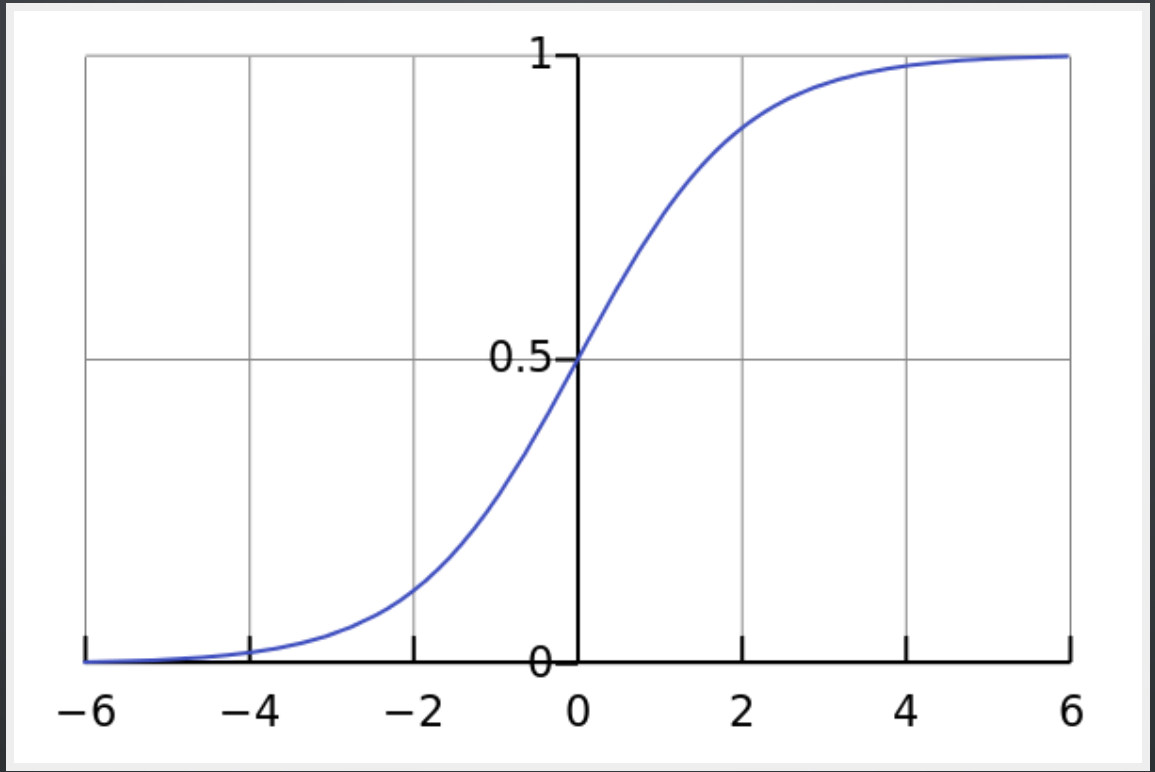
- Spam recognition
- Advertisement
- Combination of other classifiers
- Computer vision
- Calorimeter in detectors

LOGISTIC REGRESSION

$$g(x) = \langle w, x \rangle$$

Encoding probabilities with logistic function

$$p(s|x) = \sigma(g(x)), \sigma(x) = \frac{1}{1 + e^{-x}}$$



DIFFERENT REGULARIZATIONS

Initially we optimize log-likelihood

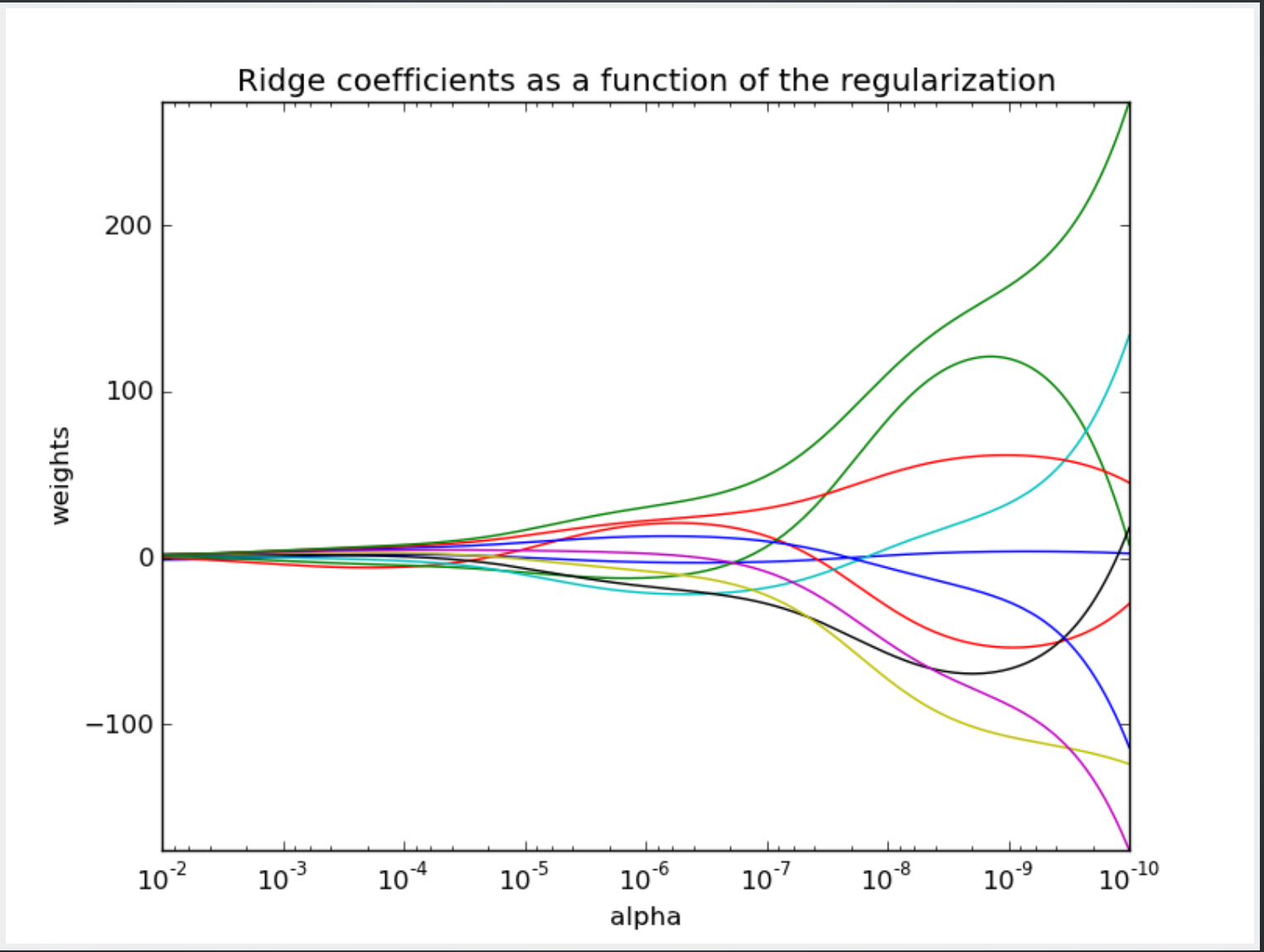
$$Q = \frac{1}{n}\sum_{i=1}^{n}\mathcal{L}(M_i) \to \min$$

$$L_2: \qquad Q_2 = Q + \alpha\sum_{j=1}^{m}|w_j|^2$$

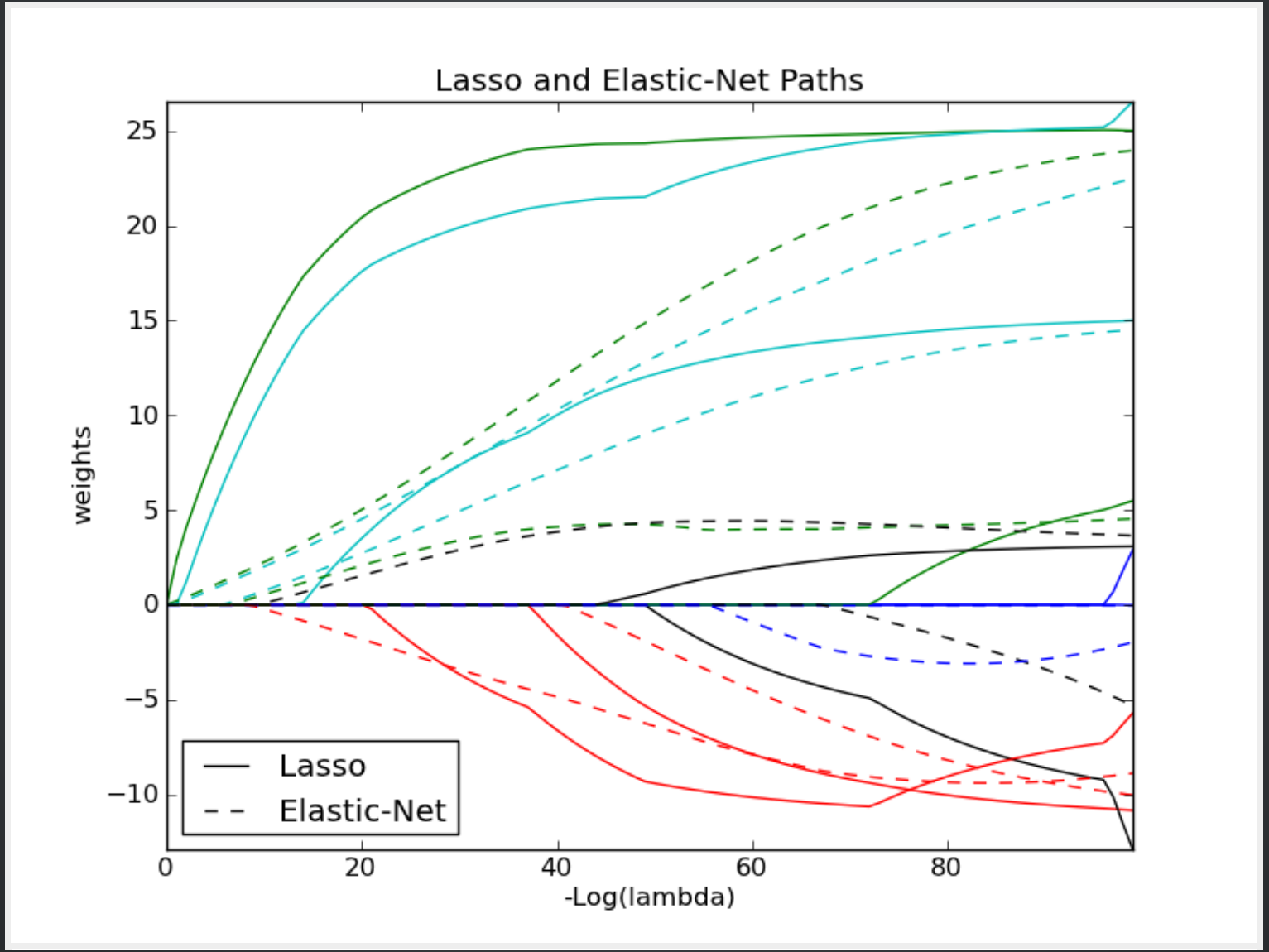
$$L_1: \qquad Q_1 = Q + \alpha\sum_{j=1}^{m}|w_j|^1$$

$$ElasticNet: Q + \alpha\sum_{j=1}^{m}|w_j|^2 + \beta\sum_{j=1}^{m}|w_j|^1$$

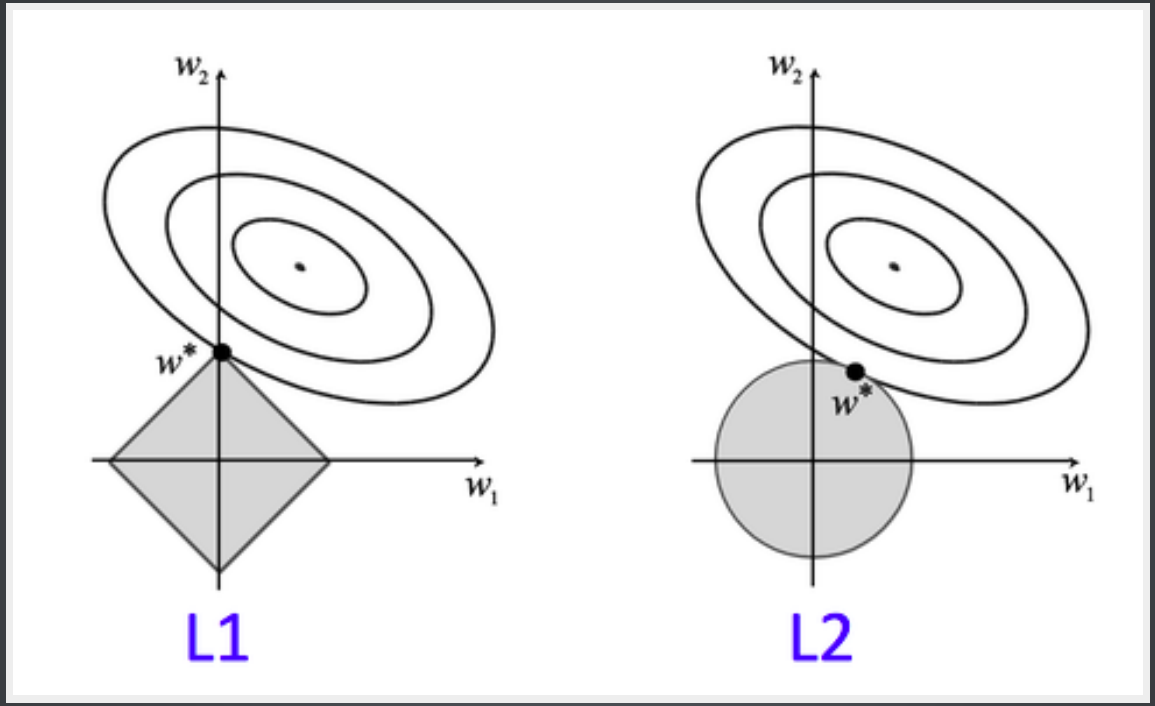
DEPENDENCE ON α (L_2)



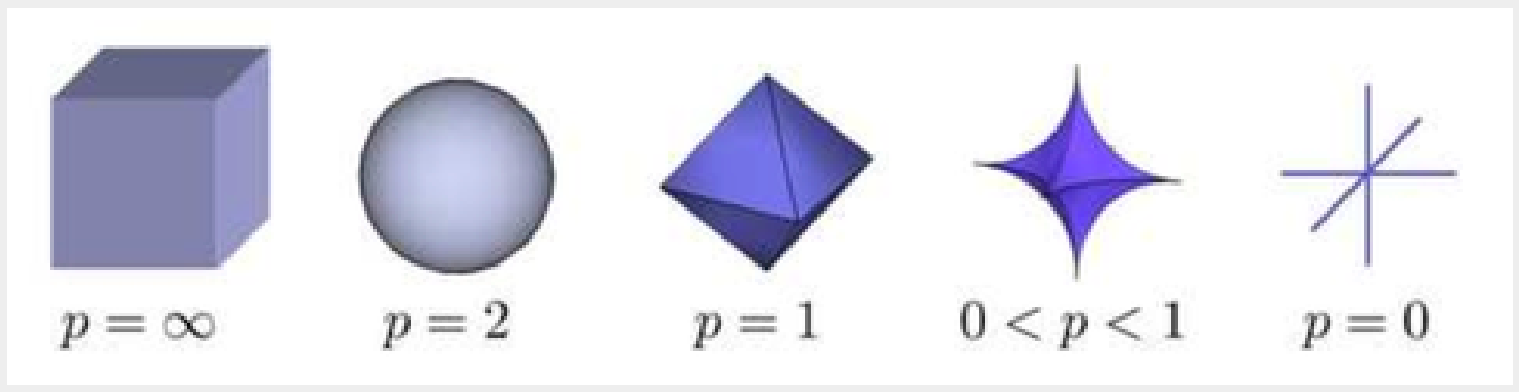
DEPENDENCE ON α (L_1, ELASTICNET)



WHY L_1 WORKS AS FEATURE SELECTOR?

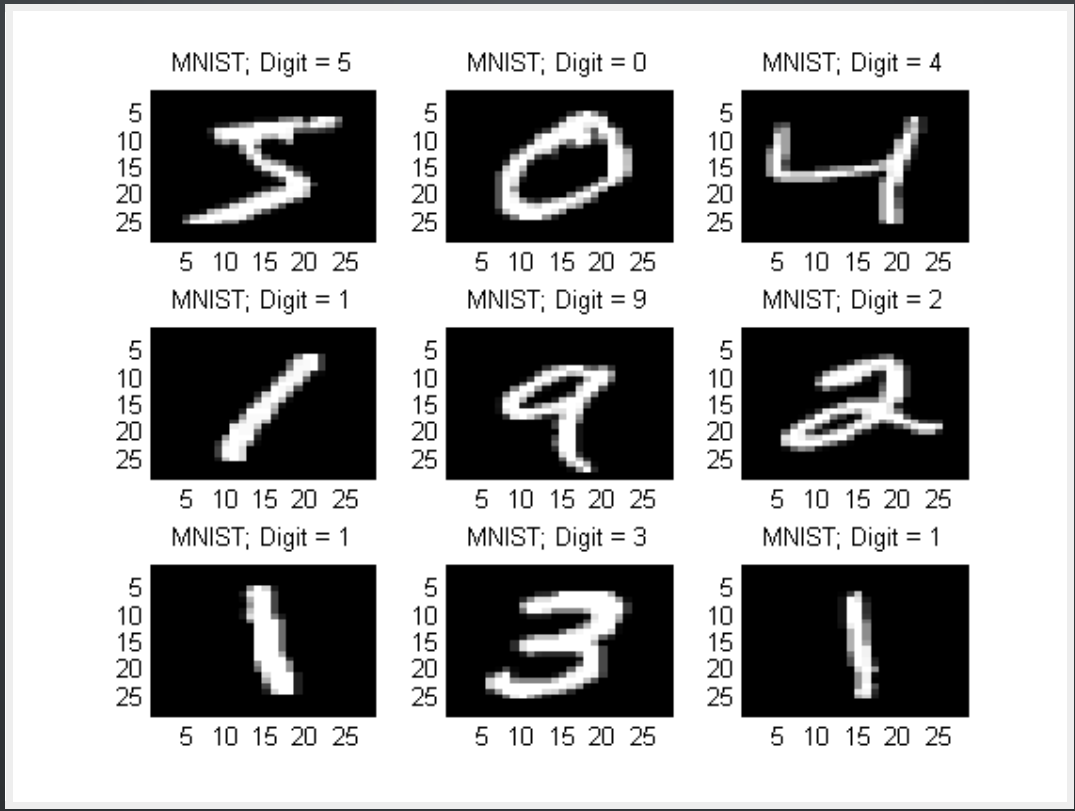


DIFFERENT NORMS

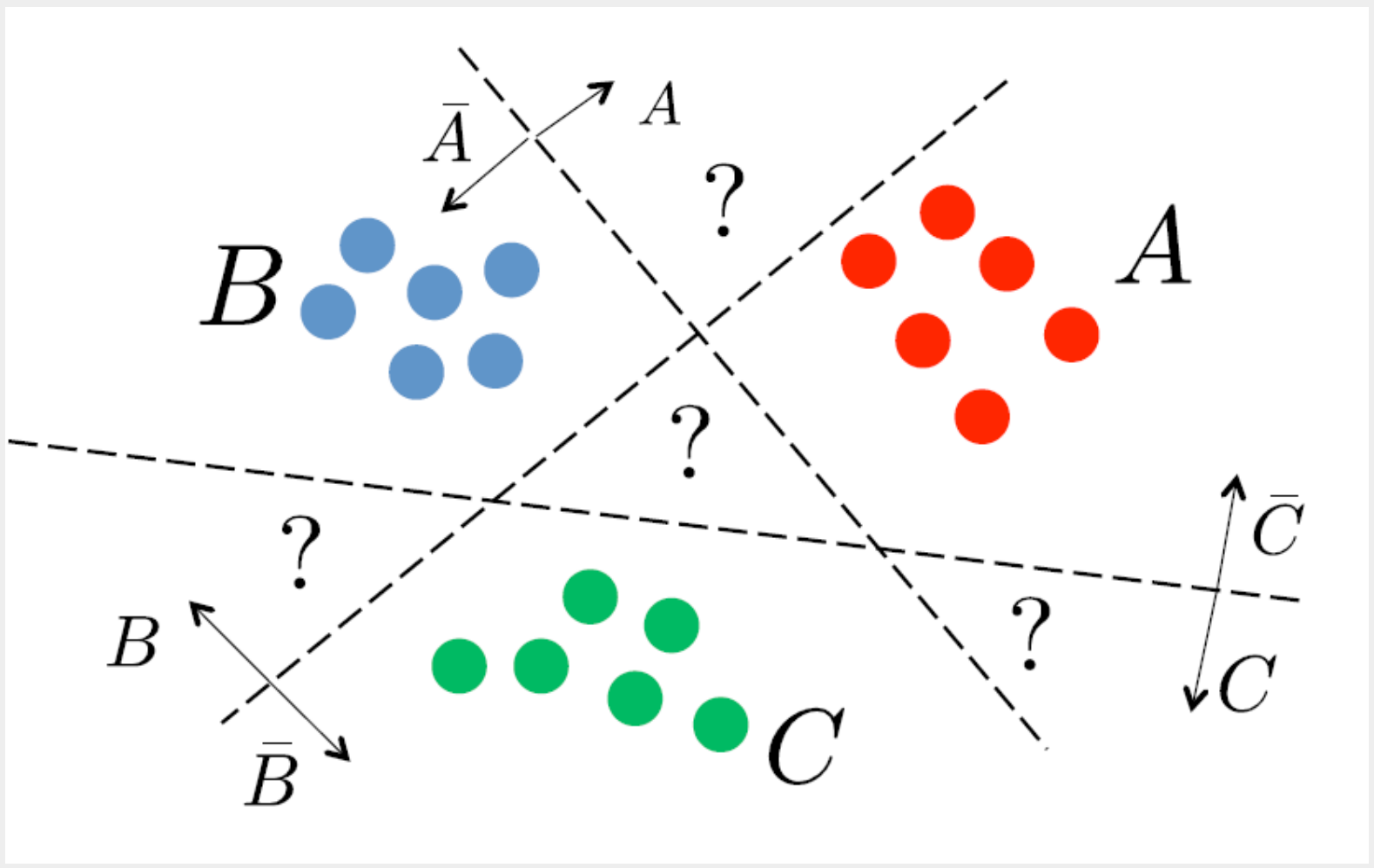


- What is L_0?
- Why nobody uses L_p with $p < 1$?

MULTICLASS CLASSIFICATION

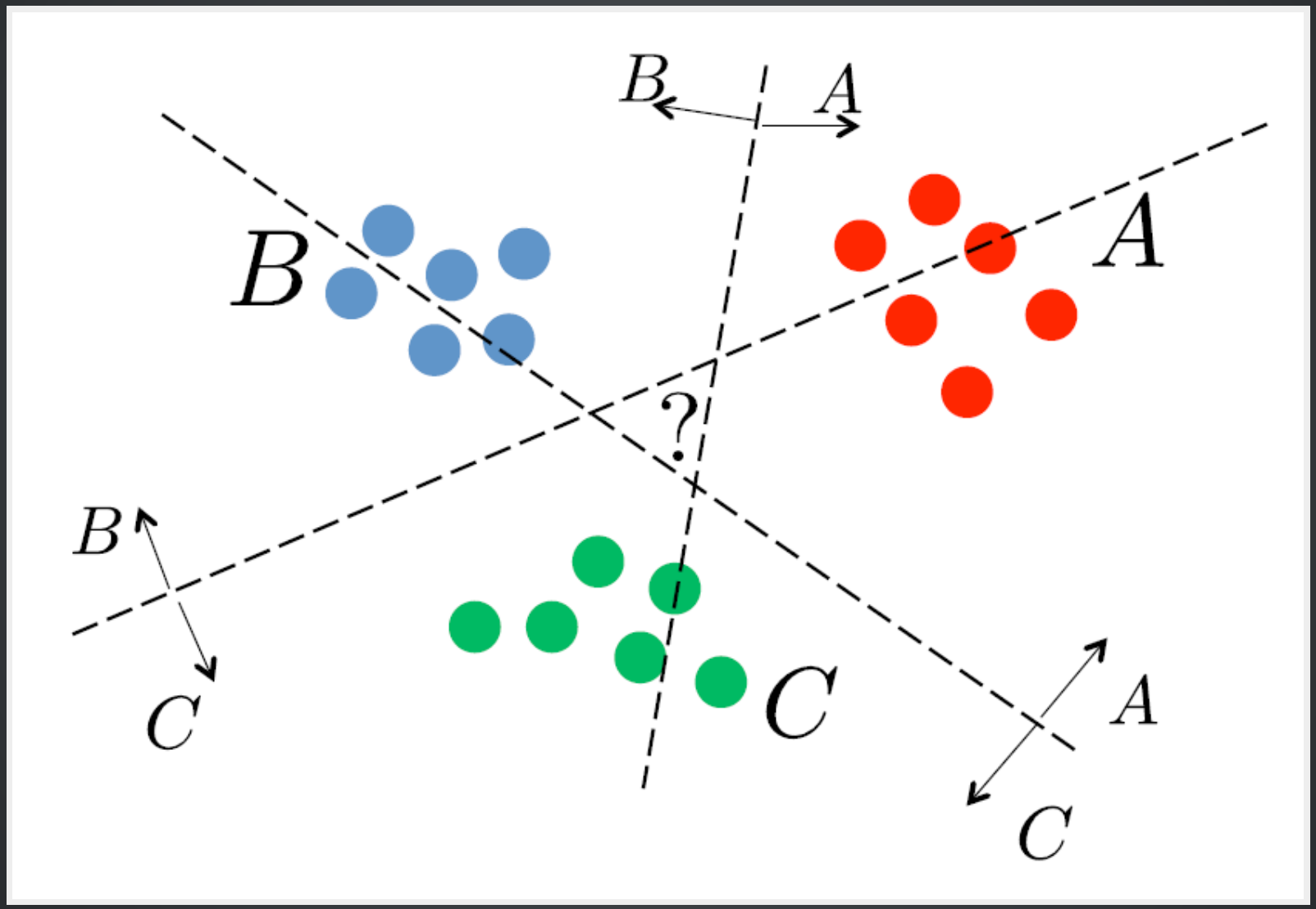


MULTICLASS CLASSIFICATION: ONE-VS-ALL

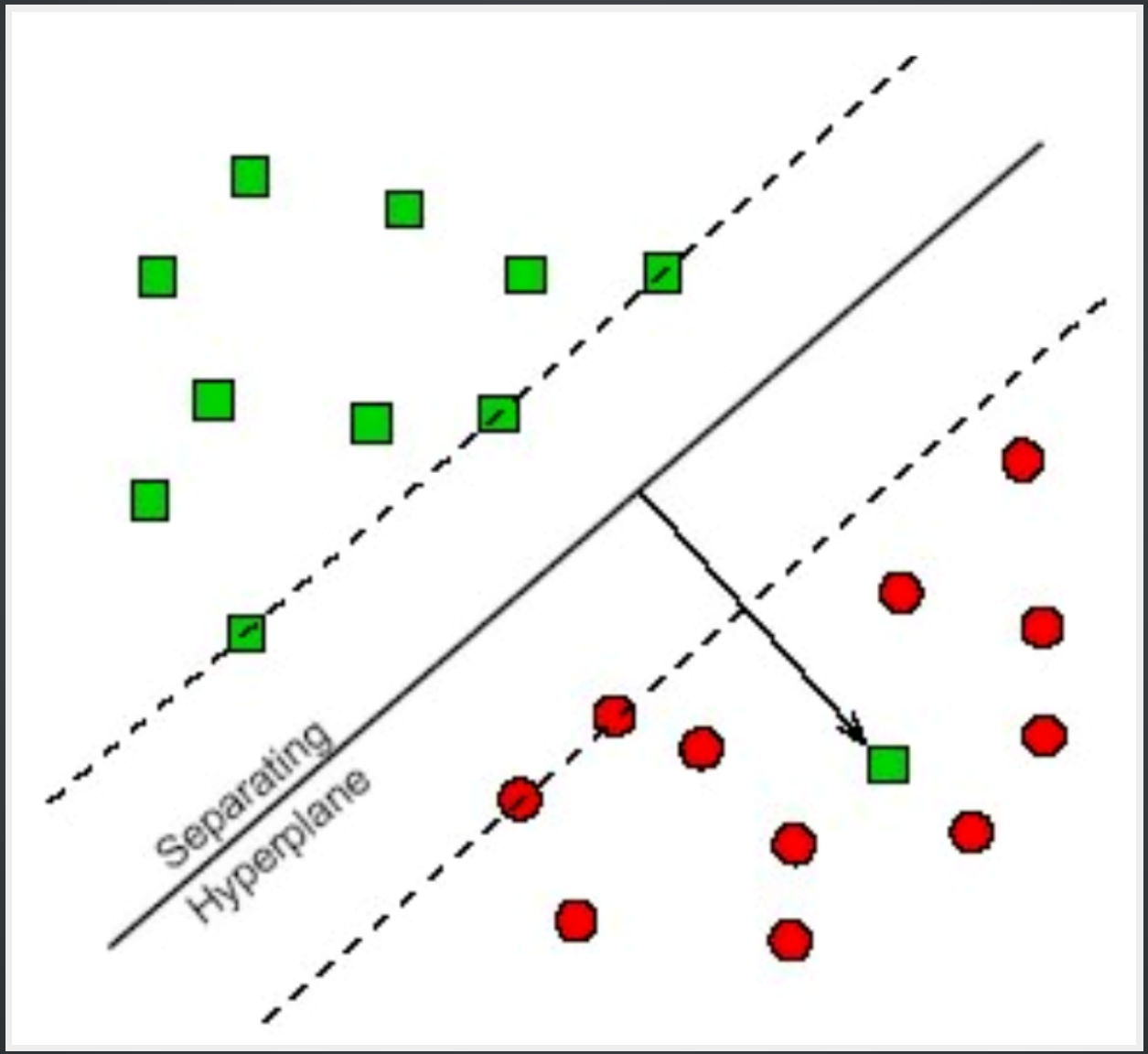


Convertation to probabilities: softmax

MULTICLASS CLASSIFICATION: ONE-VS-ONE



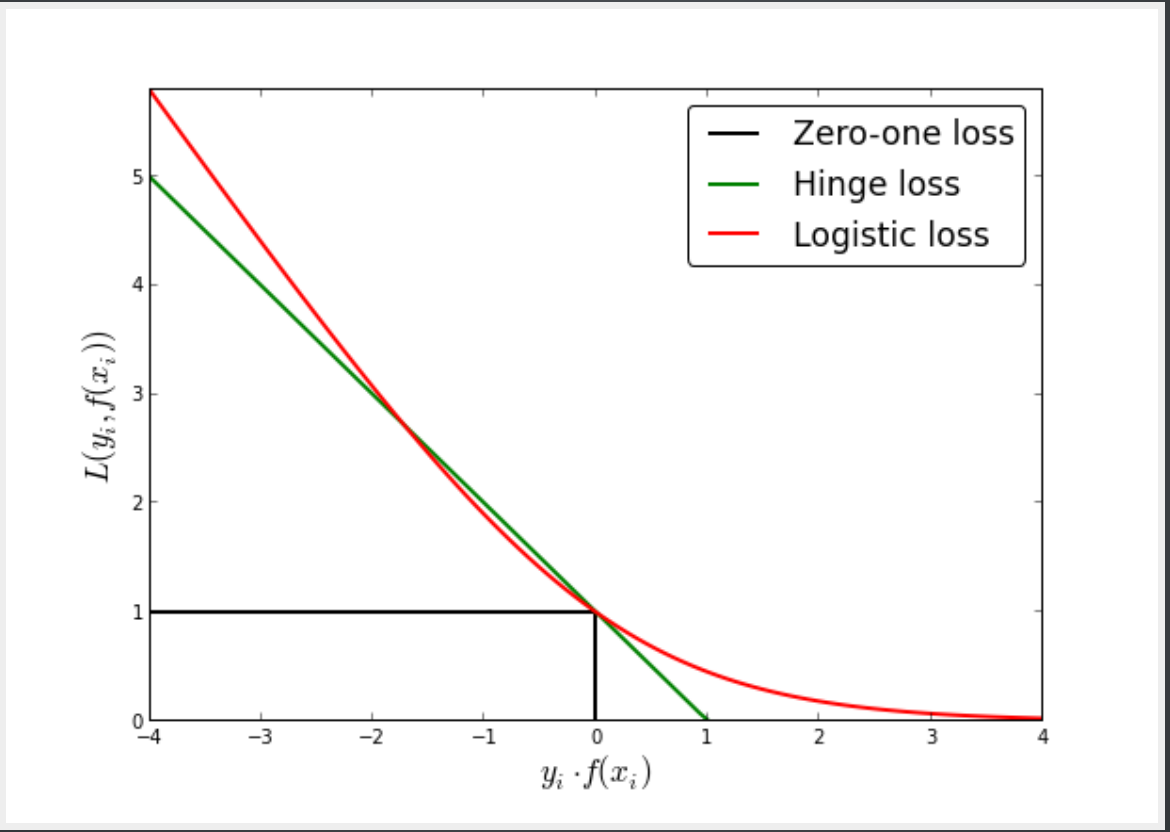
SVM



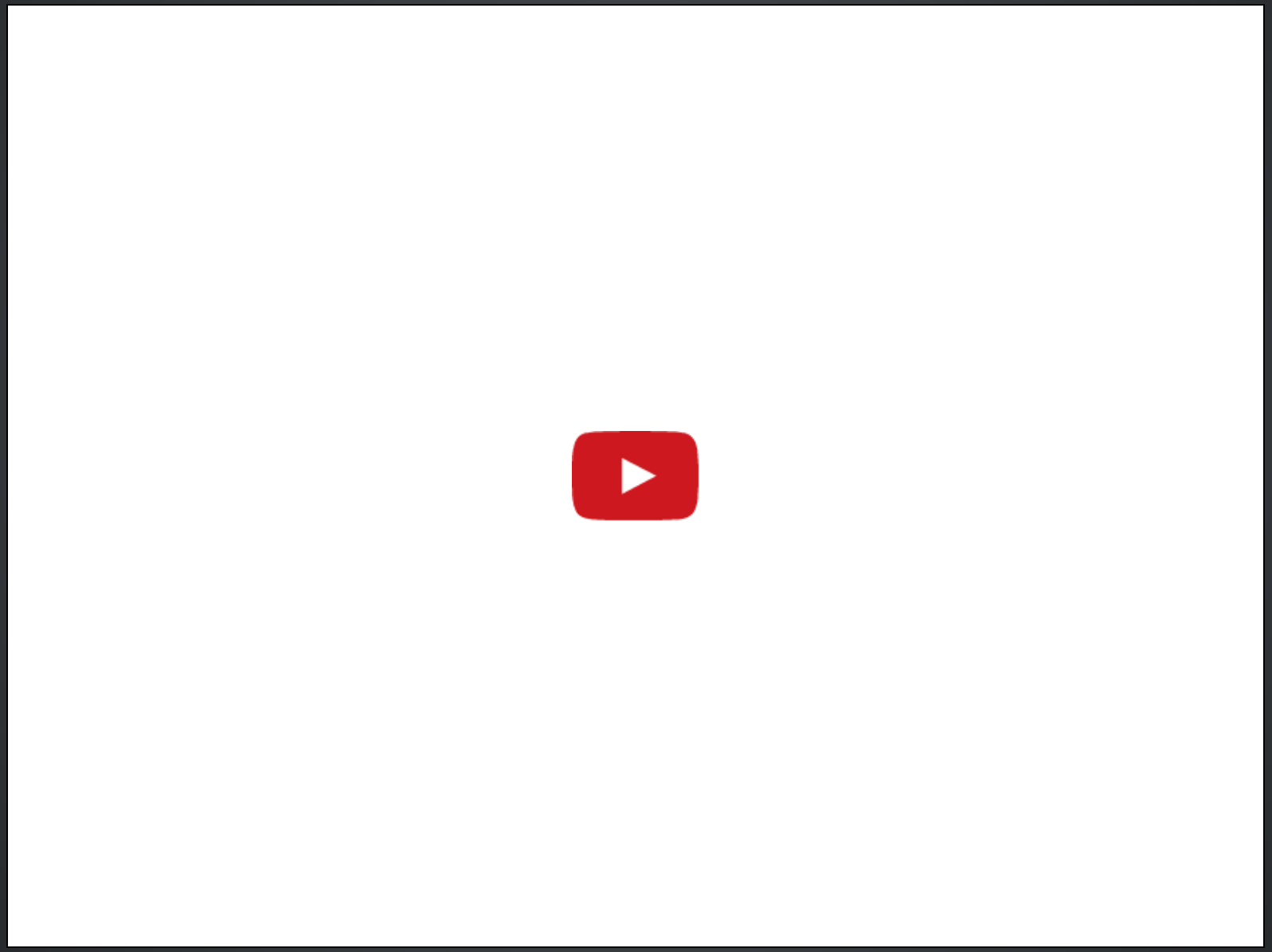
LOSS MATTER

SVM uses Hinge loss:

$$\mathcal{L}(M_i) = \max\{0, 1 - M_i\}$$



KERNEL TRICK



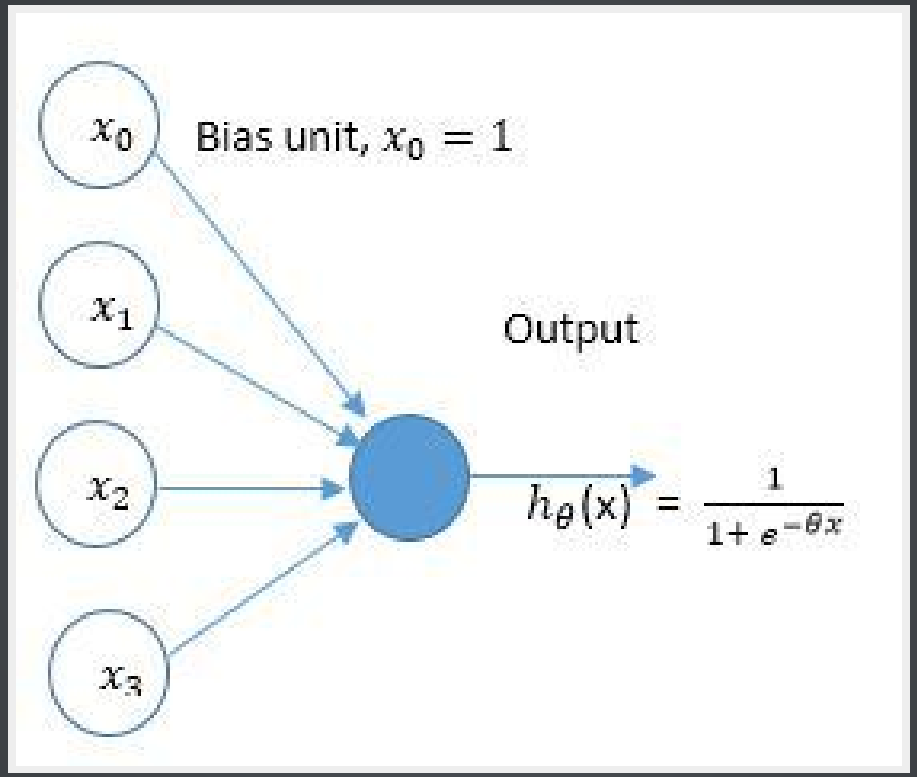
KERNEL TRICK

- Works with other linear models
- But only SVM 'selects' events (support vectors)
- How to verify the quality of trained svm fast?

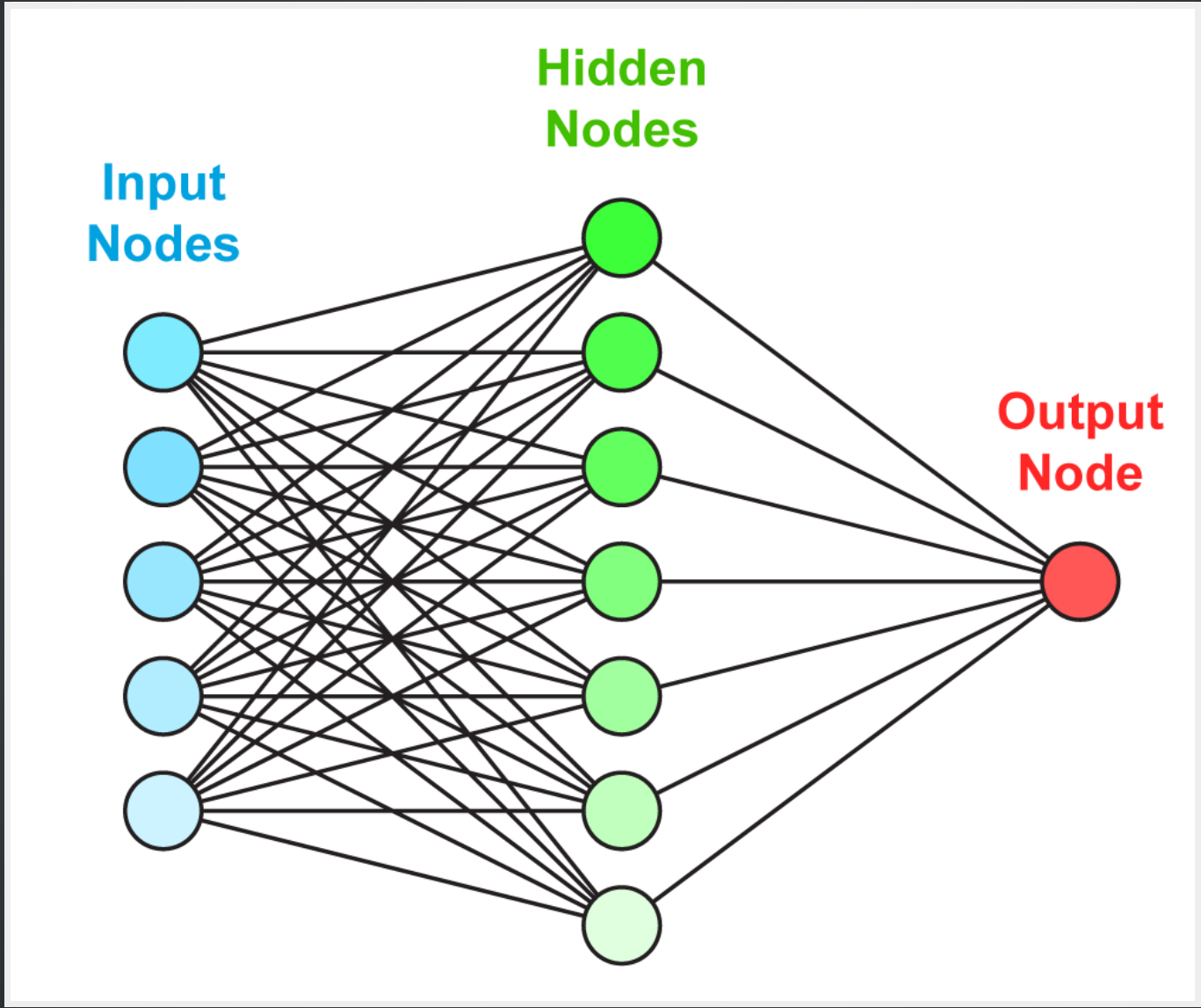
NEURAL NETWORKS

LOGISTIC REGRESSION

is the simplest neural network :)



2-LAYER NEURAL NETWORK



ANN DEMONSTRATION

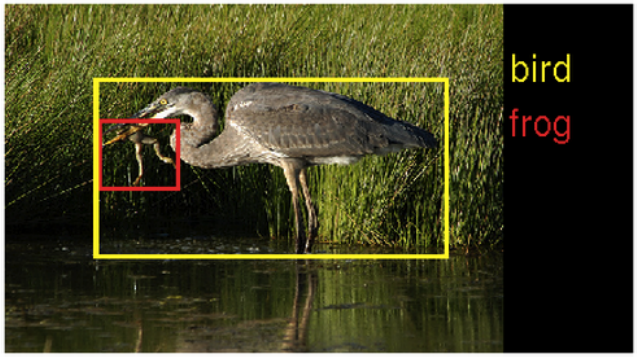
- Long training, fast prediction
- Regularization
- Many training algorithms
- Backpropagation to compute derivatives

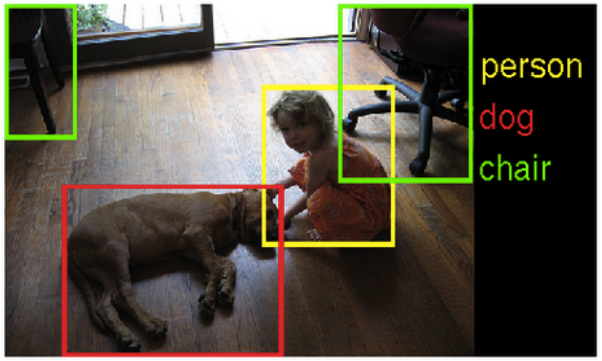
PROBLEMS WITH DEEP NETWORKS

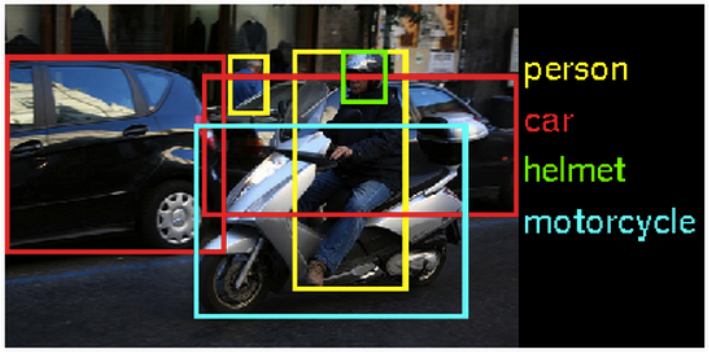
- gradient diminishes on first layers
- overfitting

IMAGES RECOGNITION PROBLEM

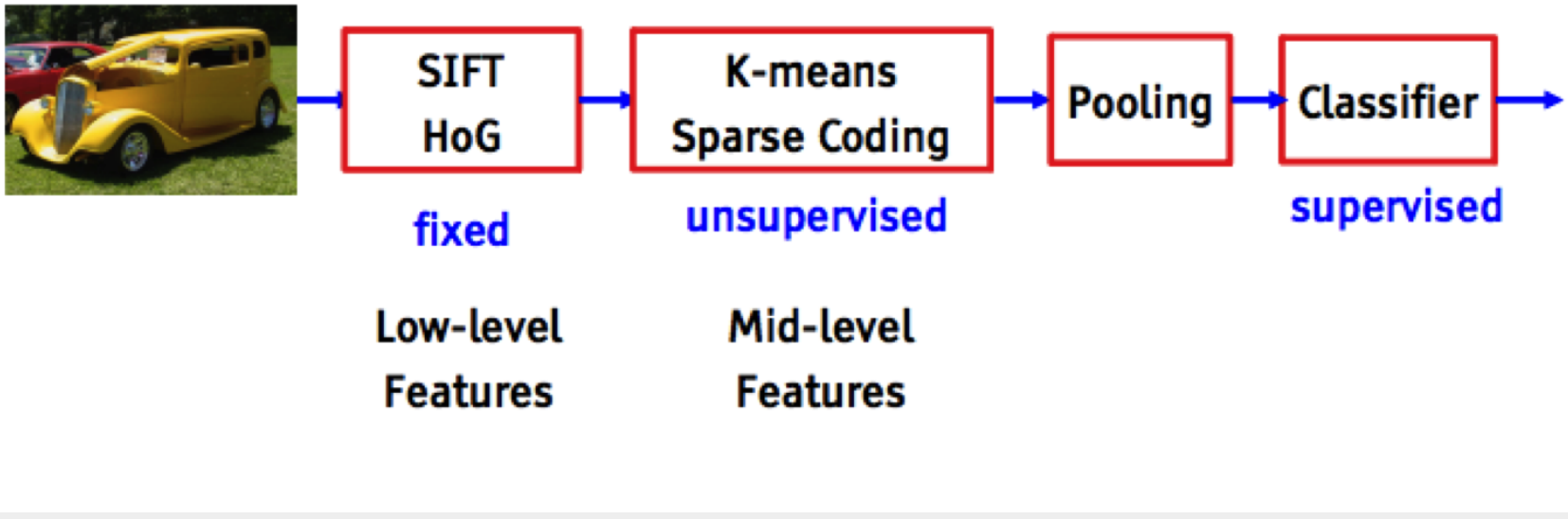
Example ILSVRC2014 images:



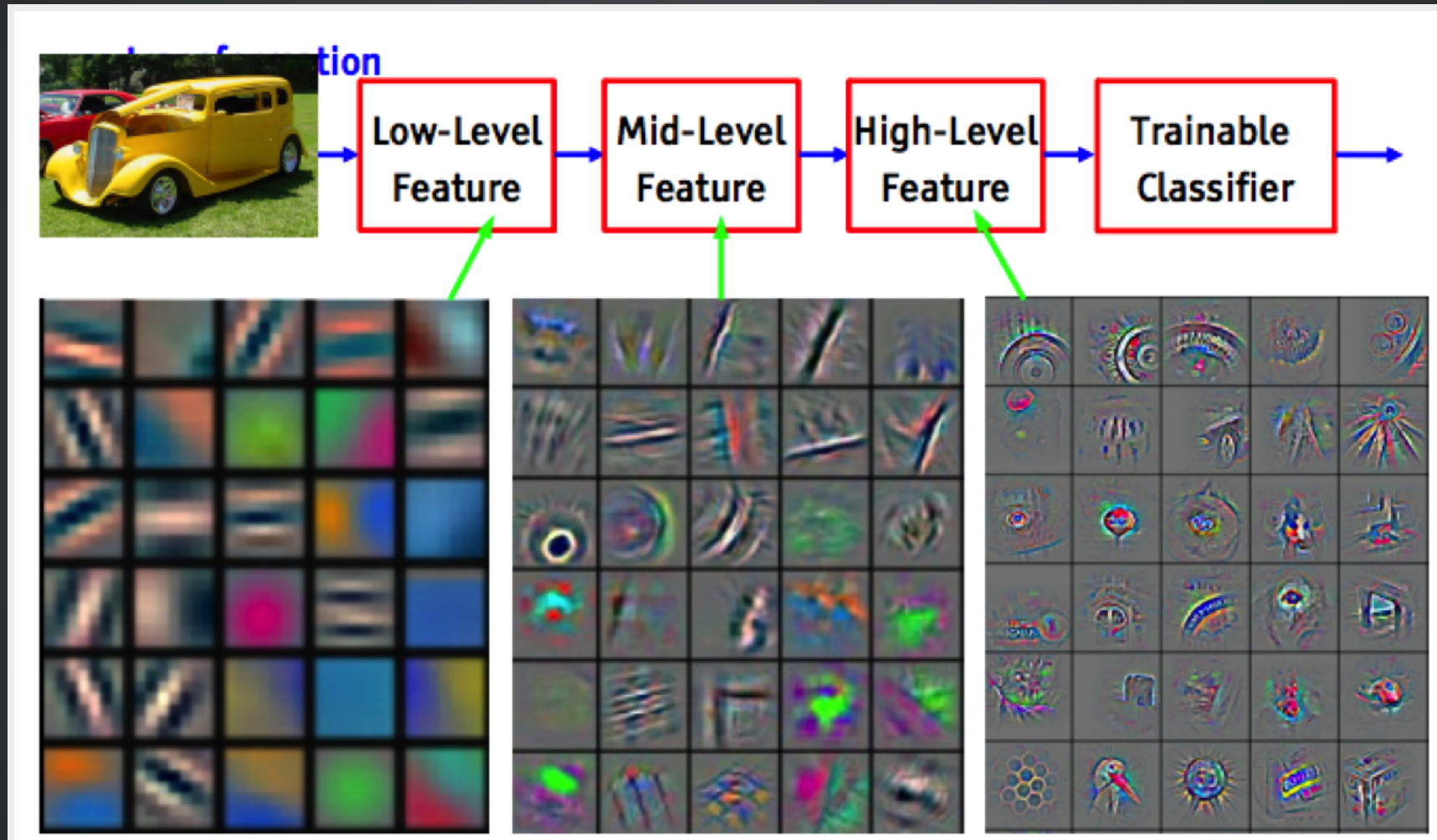




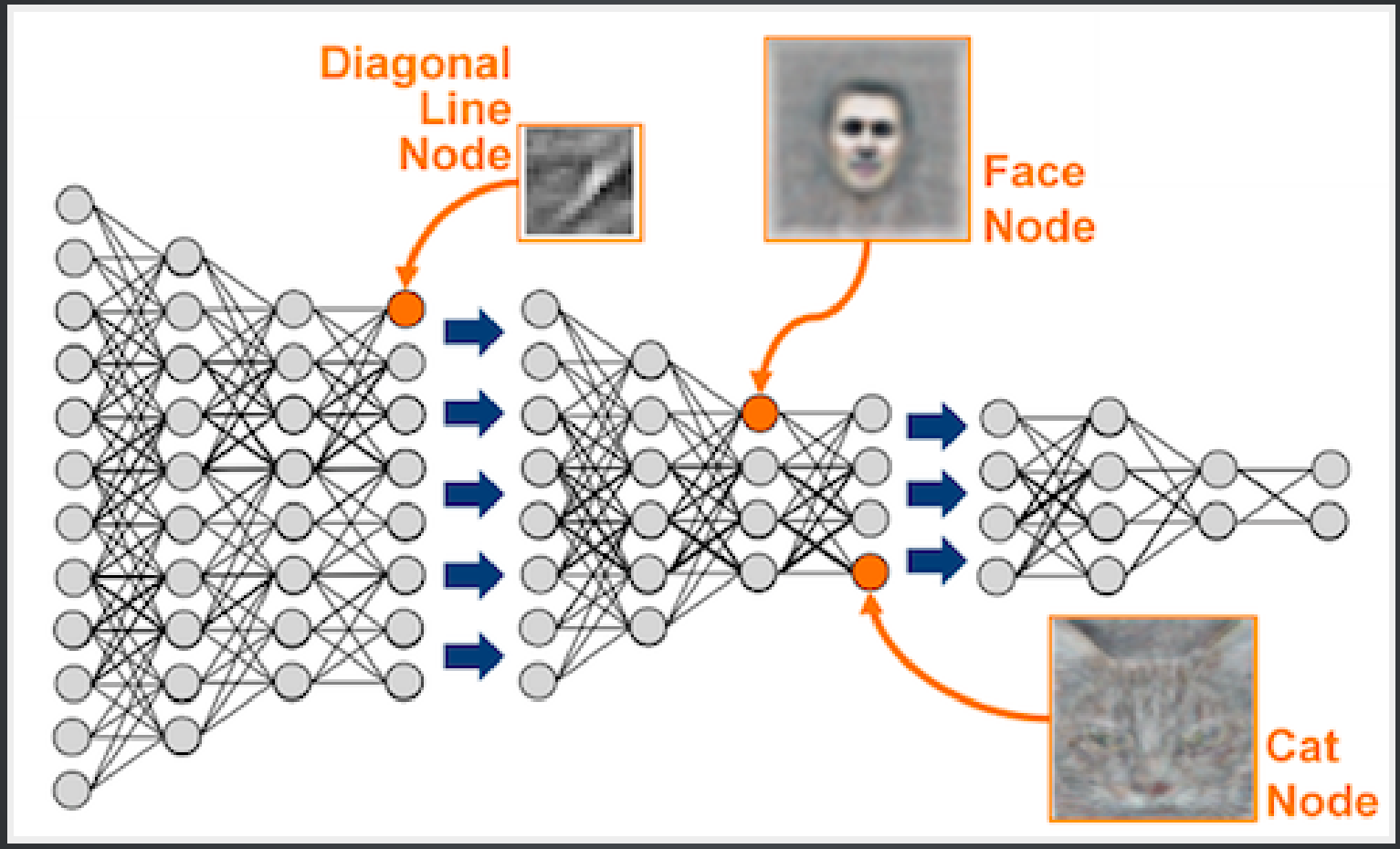
COMMON APPROACH IN IMAGE RECOGITION



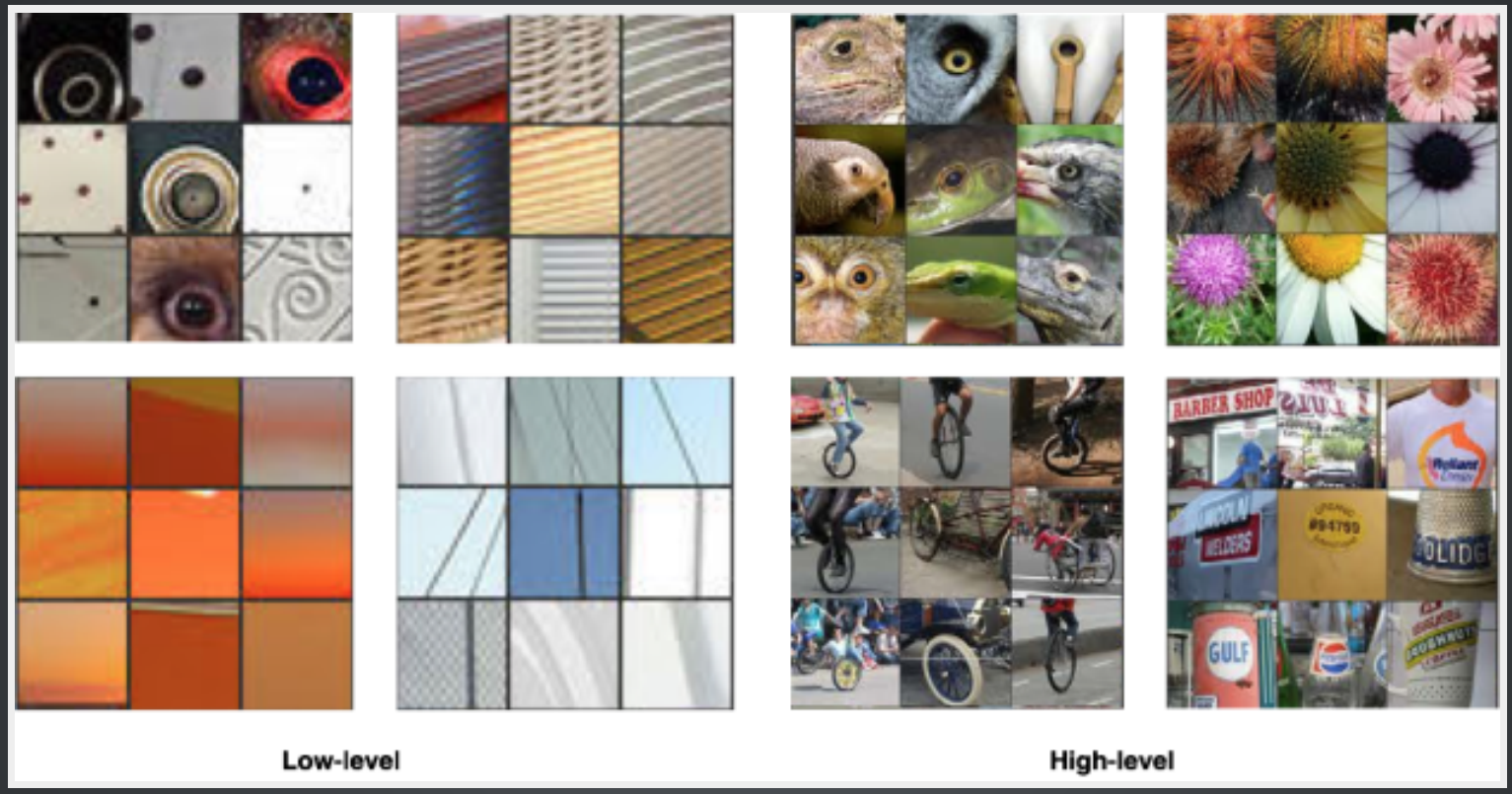
DEEP LEARNING APPROACH



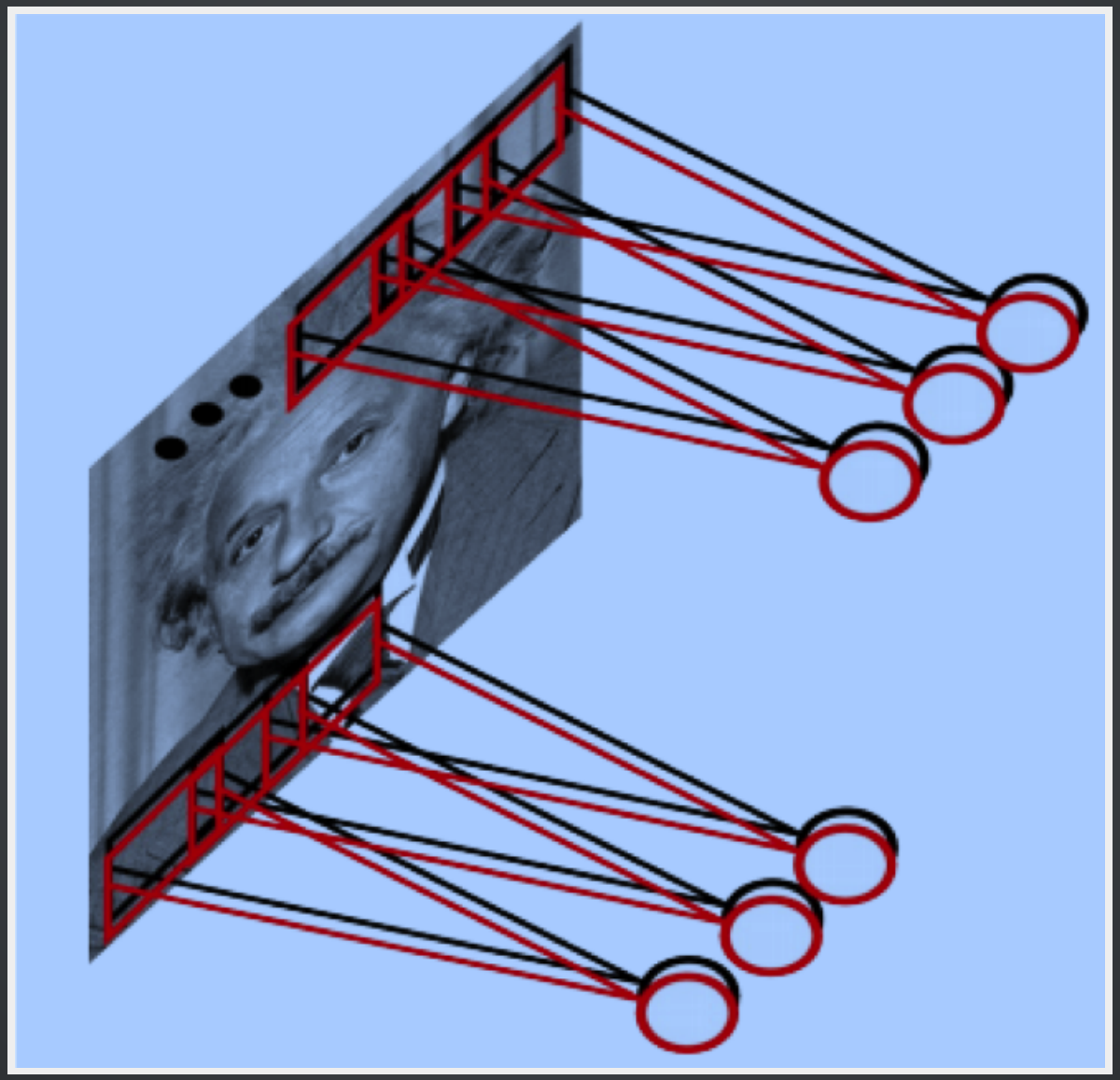
DEEP LEARNING APPROACH



DEEP LEARNING: VISUALIZATION OF PATCHES



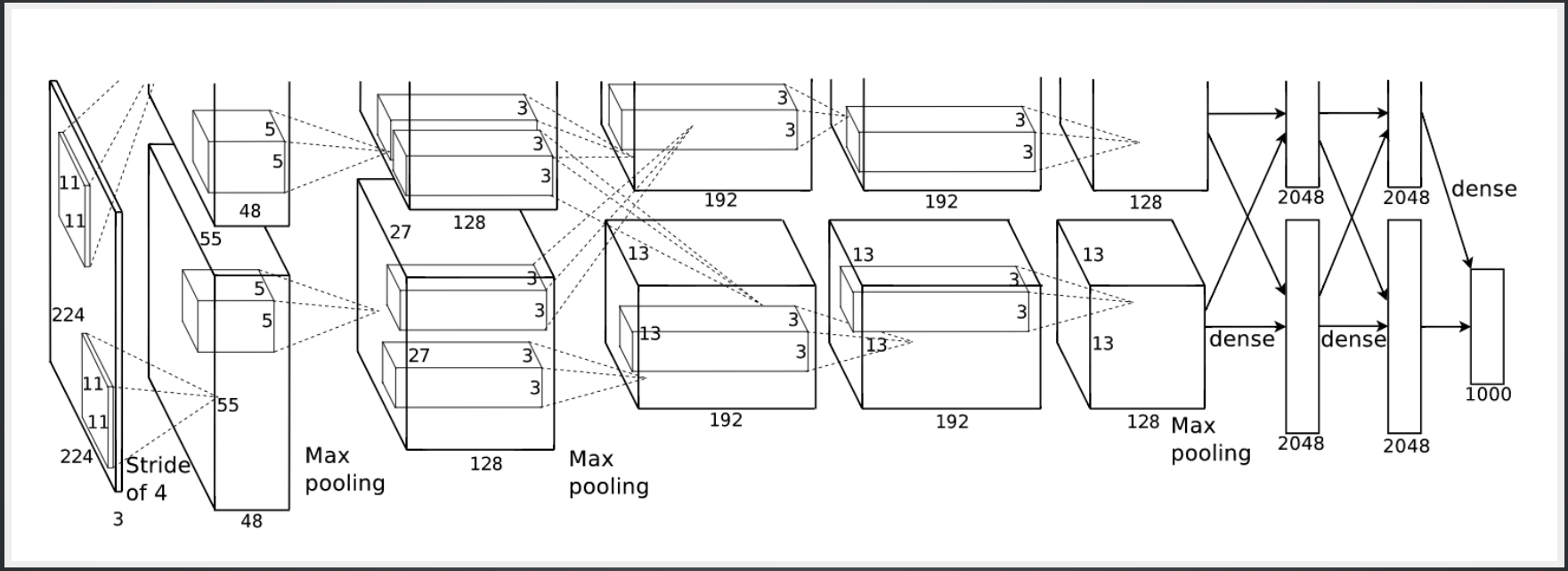
CONVOLUTIONAL NEURAL NETWORKS



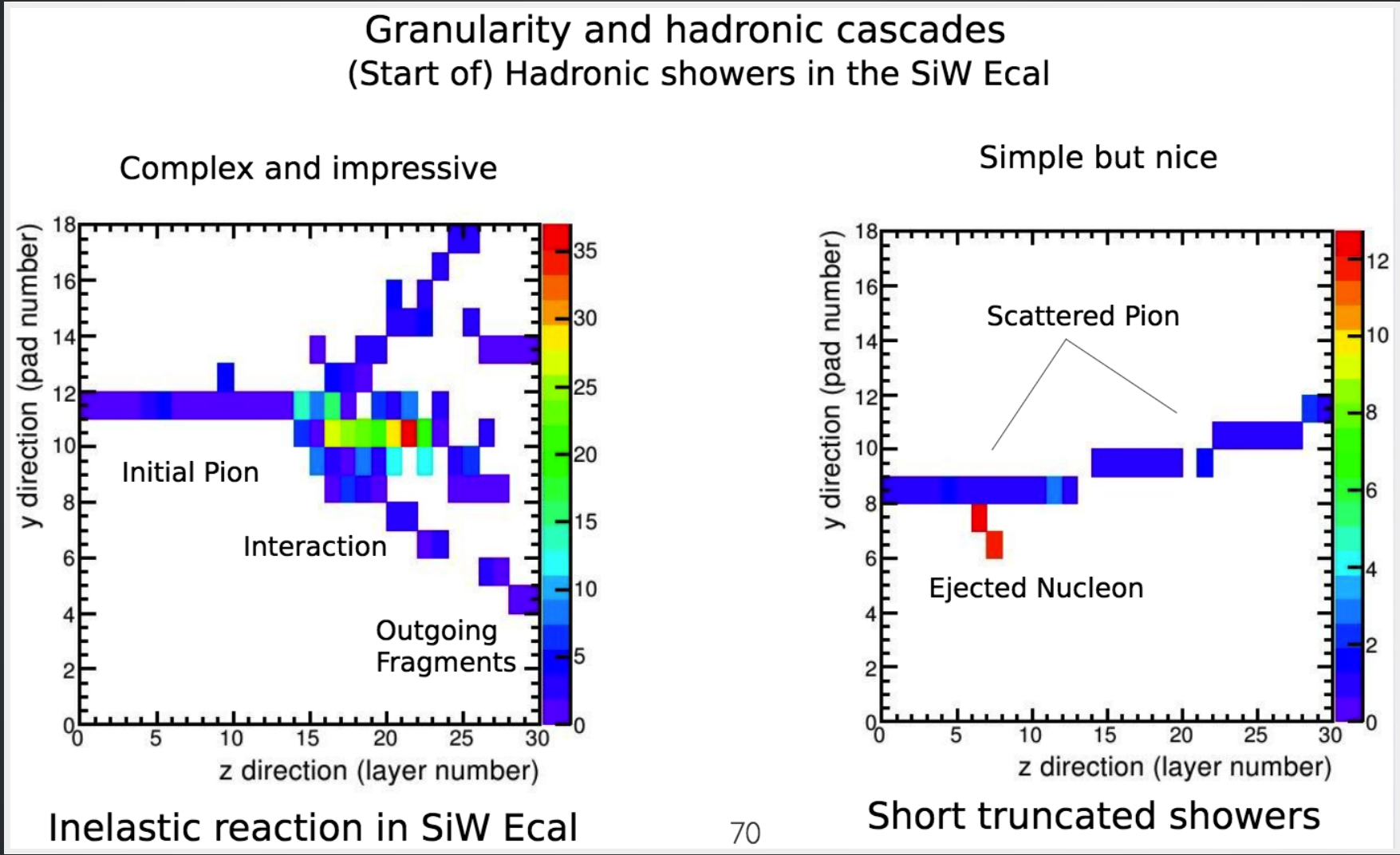
NOT SO SIMPLE

Best solution for ImageNet classification 2013

http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf



DEEP LEARNING FOR IMAGING CALORIMETERS



https://indico.cern.ch/event/316800/material/slides/0.pdf


THE END