# Machine Learning

**Sergei Gleyzer**

## PART III

**TAE 2017 Lectures**

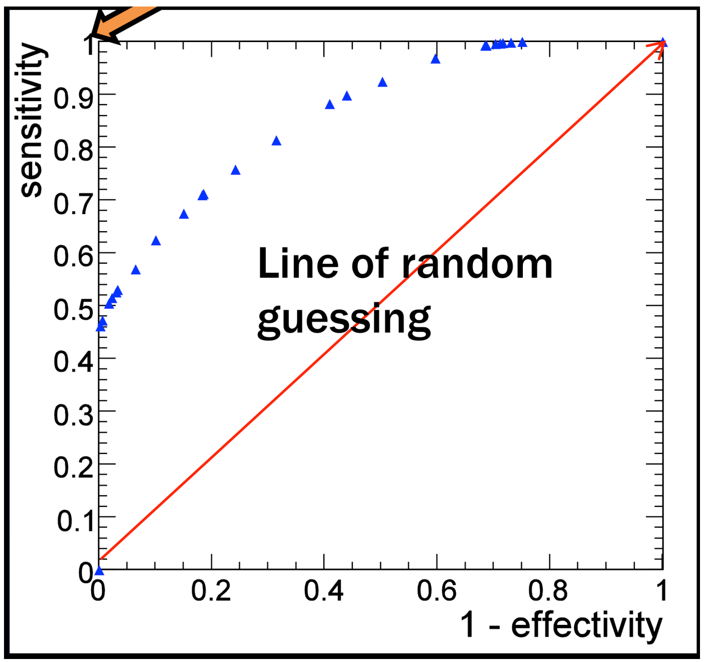
**Sep. 6, 2017**

# Classifier Performance

# Classifier Performance

## Receiver Operating Characteristic (ROC)

Perfect Classifier




**Commonly used metric**

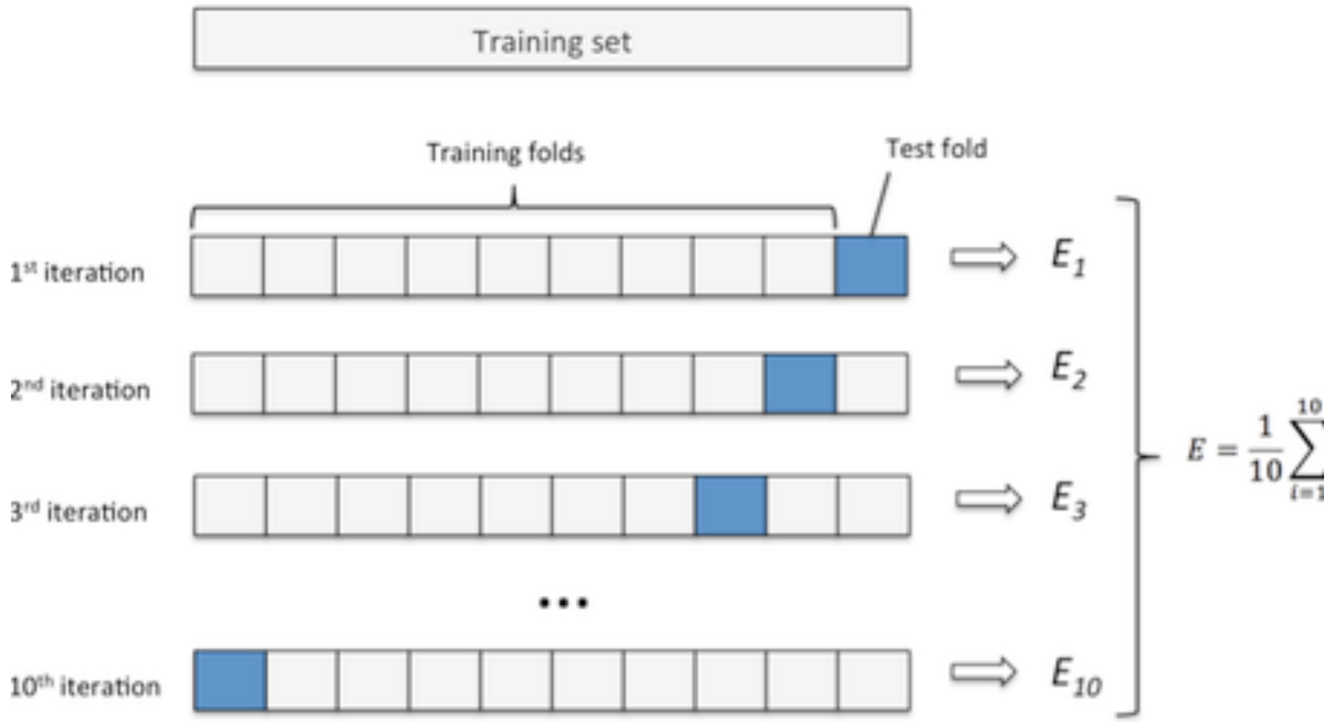
Shows the **relationship** between correctly classified positive cases (sensitivity) and incorrectly classified negative cases (1-effectivity)

# Cross Validation

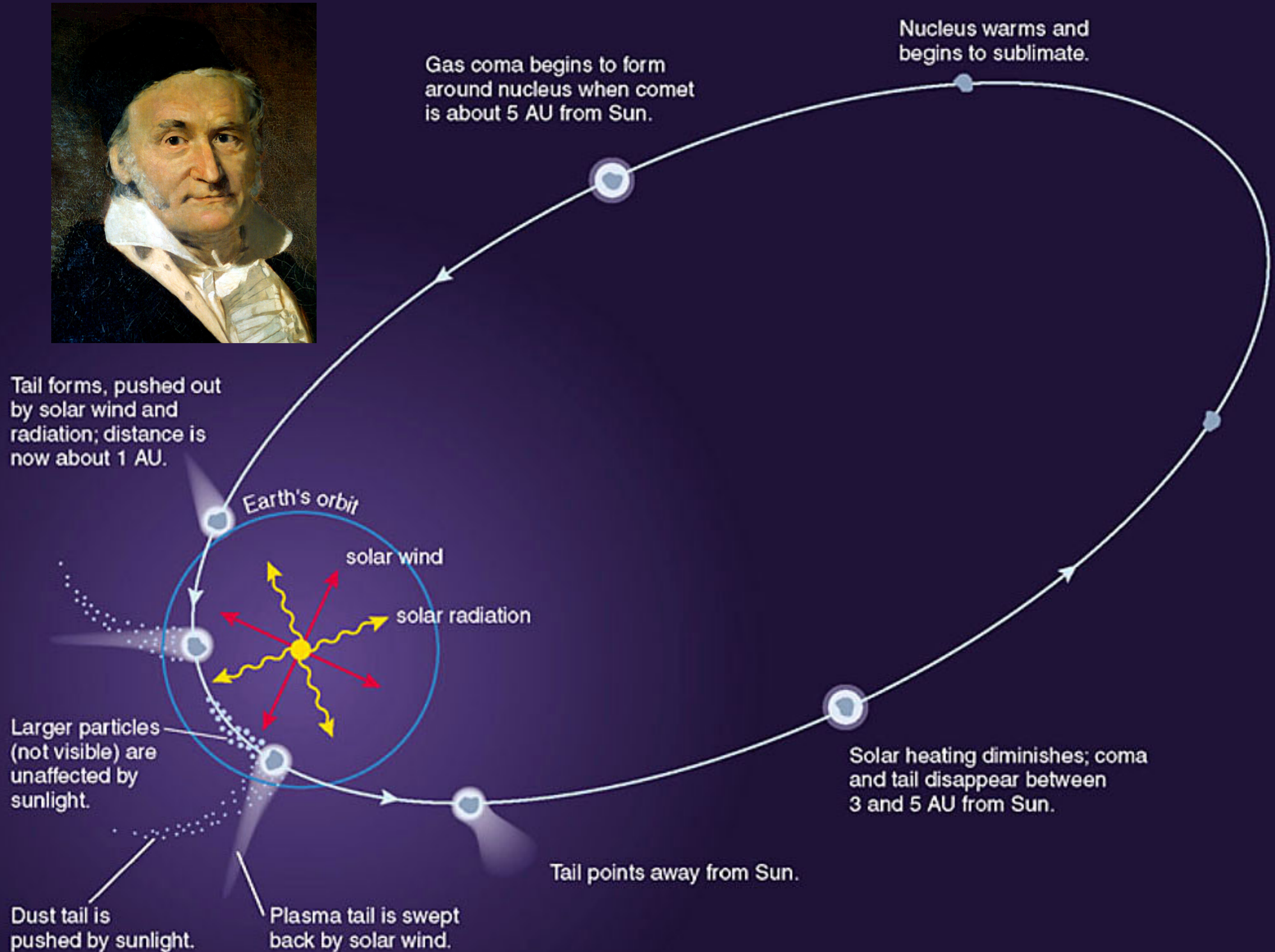
**Generalization** of train-test split for more accurate evaluation of classifier performance

- Randomly split dataset into K equal partitions
- In each fold use K-1 samples to train, leftover to test

# Cross Validation

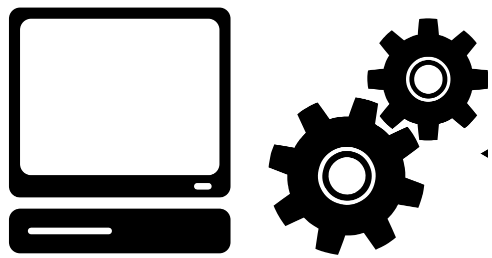


# Function Estimation


Nucleus warms and begins to sublimate.
Gas coma begins to form around nucleus when comet is about 5 AU from Sun.
Tail forms, pushed out by solar wind and radiation; distance is now about 1 AU.
Earth's orbit
solar wind
solar radiation
Larger particles (not visible) are unaffected by sunlight.
Dust tail is pushed by sunlight.
Plasma tail is swept back by solar wind.
Tail points away from Sun.
Solar heating diminishes; coma and tail disappear between 3 and 5 AU from Sun.

# Regression

## Modify evaluation in induction algorithm



**Maximum separation** → **Minimal variance**

# Photon Energy

## Inputs:

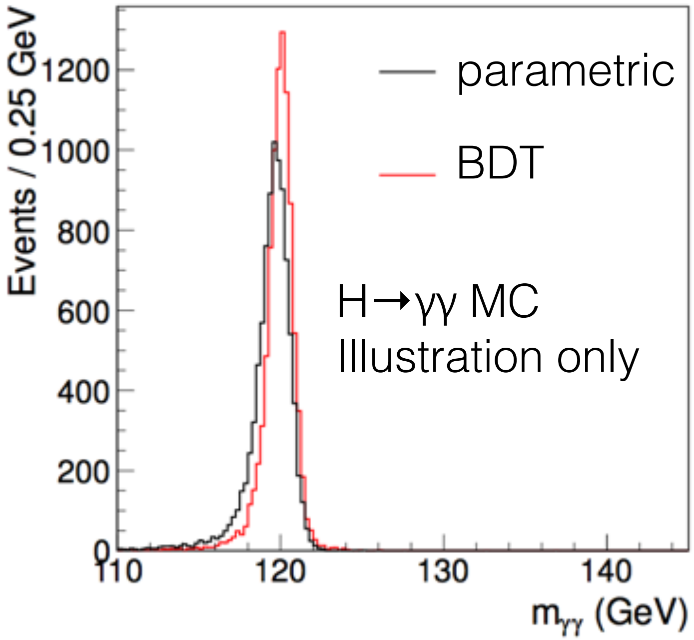
photon coordinates

photon shower information

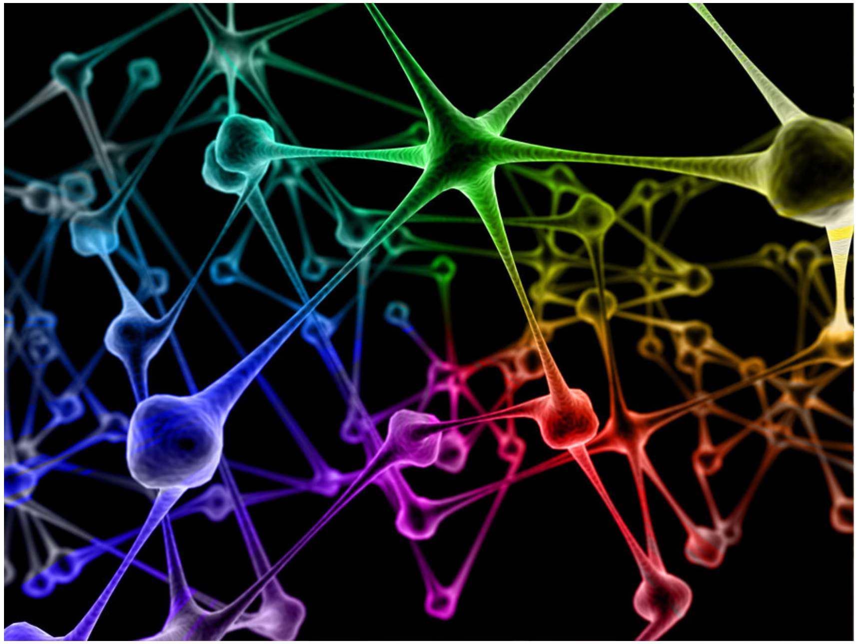
median event energy

## Target Output:

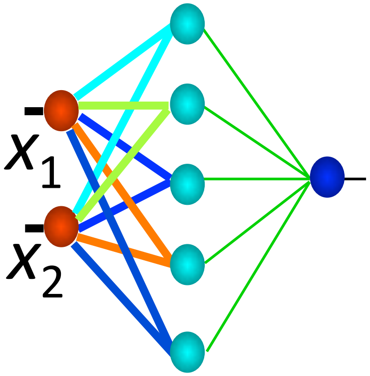
$E_{MEASURED}/E_{TRUE}$

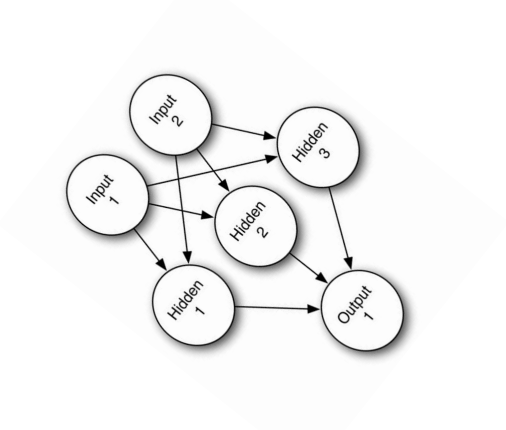



**10-30% improvement with shallow ML**

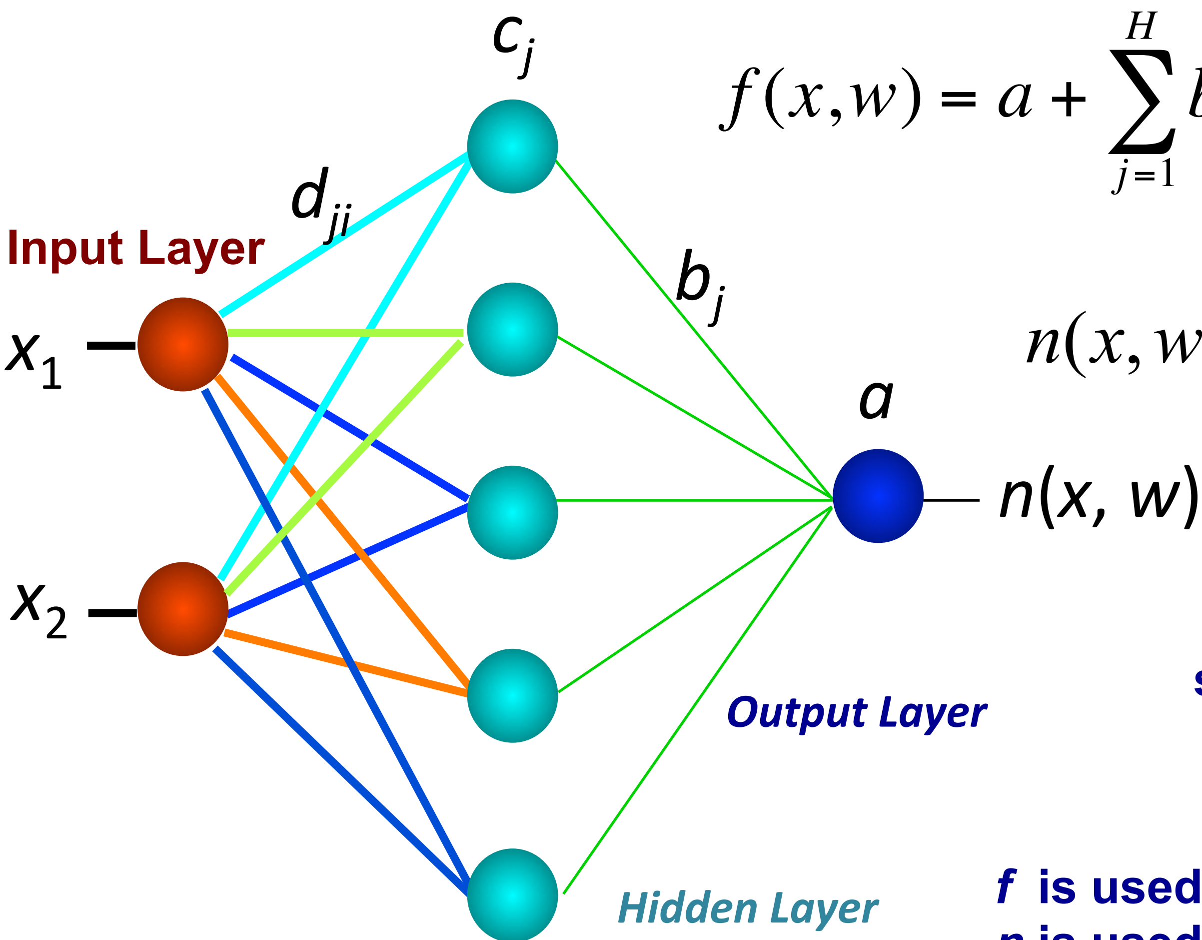


# Artificial Neural Networks



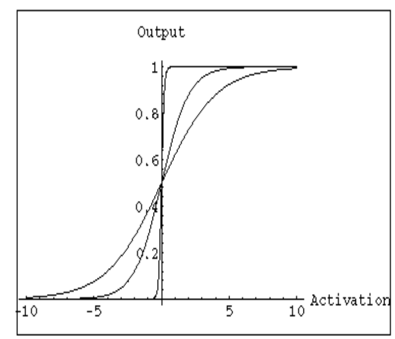
# Graphical Representation




$$f(x,w) = a + \sum_{j=1}^{H} b_j \tanh\left[c_j + \sum_{i=1}^{I} d_{ji} x_i\right]$$

$$n(x,w) = \frac{1}{1 + \exp[-f(x,w)]}$$

sigmoid

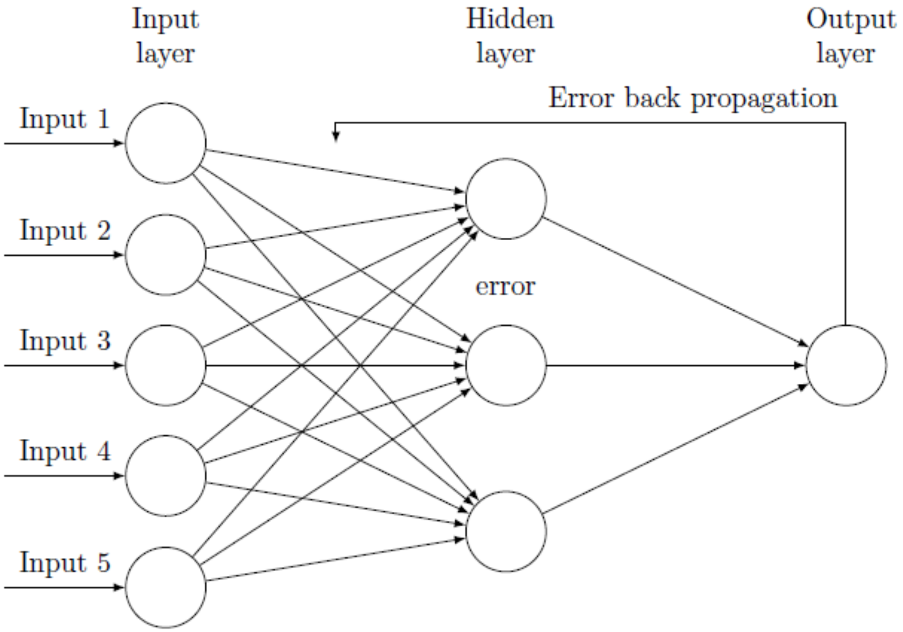



***f*** **is used for regression**
***n*** **is used for classification**
***w = a, b, c, d***

# Adjustable Weights

## Compute network weights with

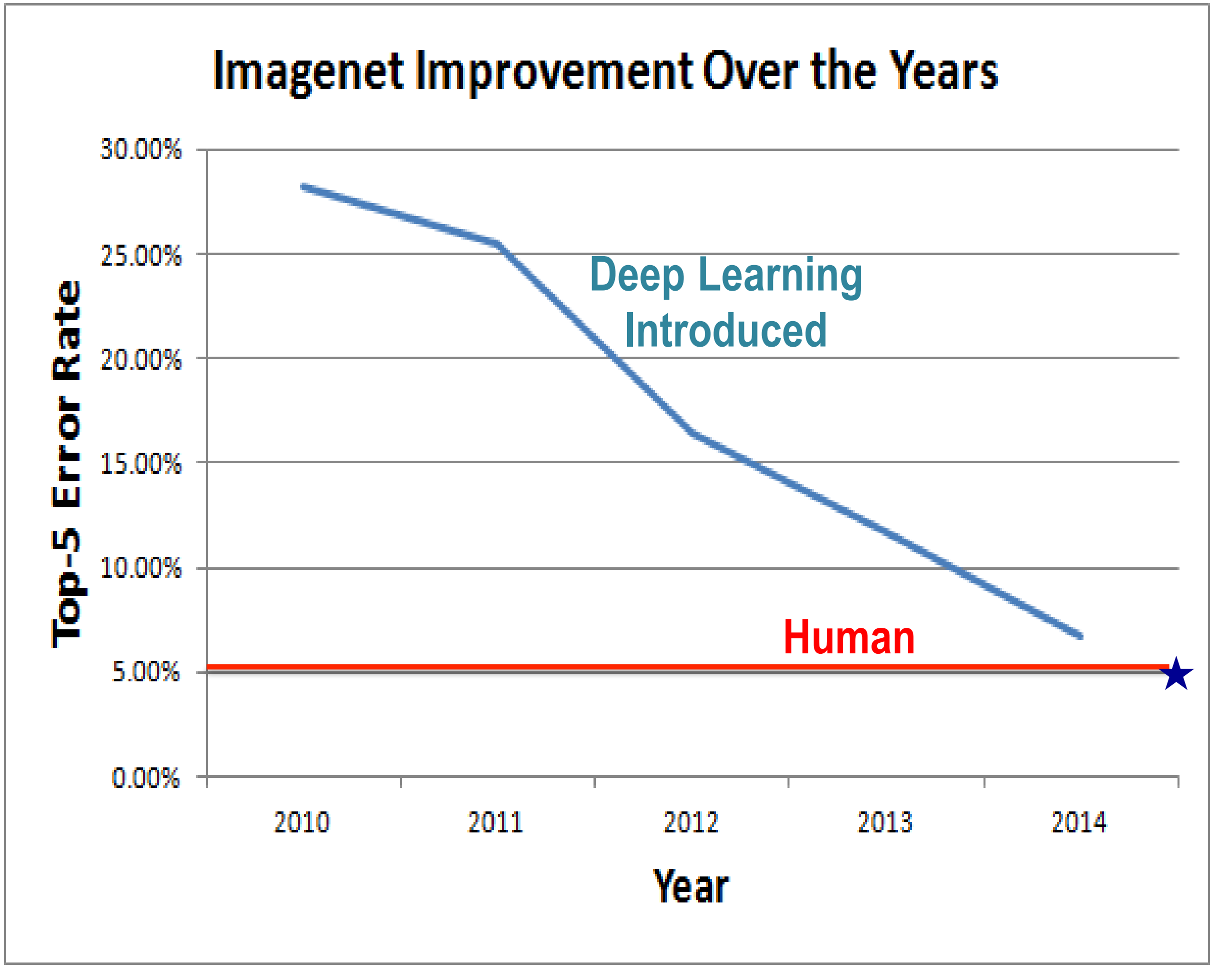
- Error gradients




**Inputs forward**

**Errors go backward!**

# Deep Learning



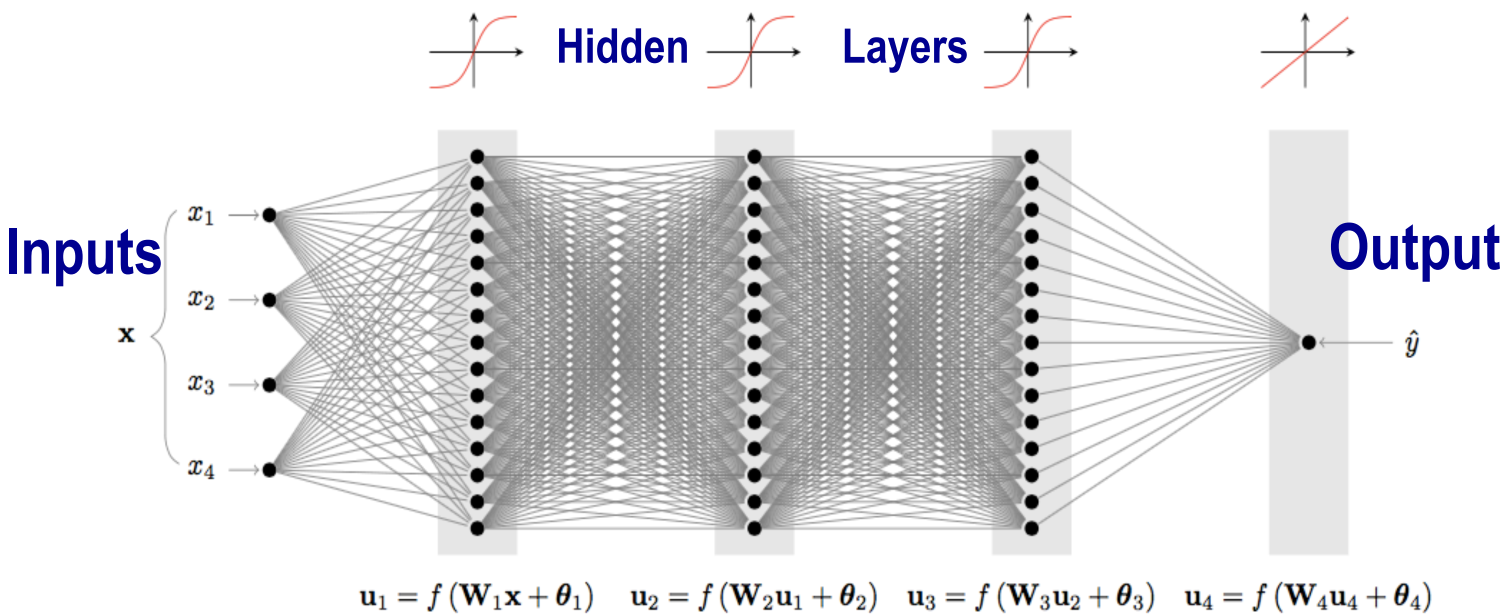

First super-human result in 2015

Google 4.9%

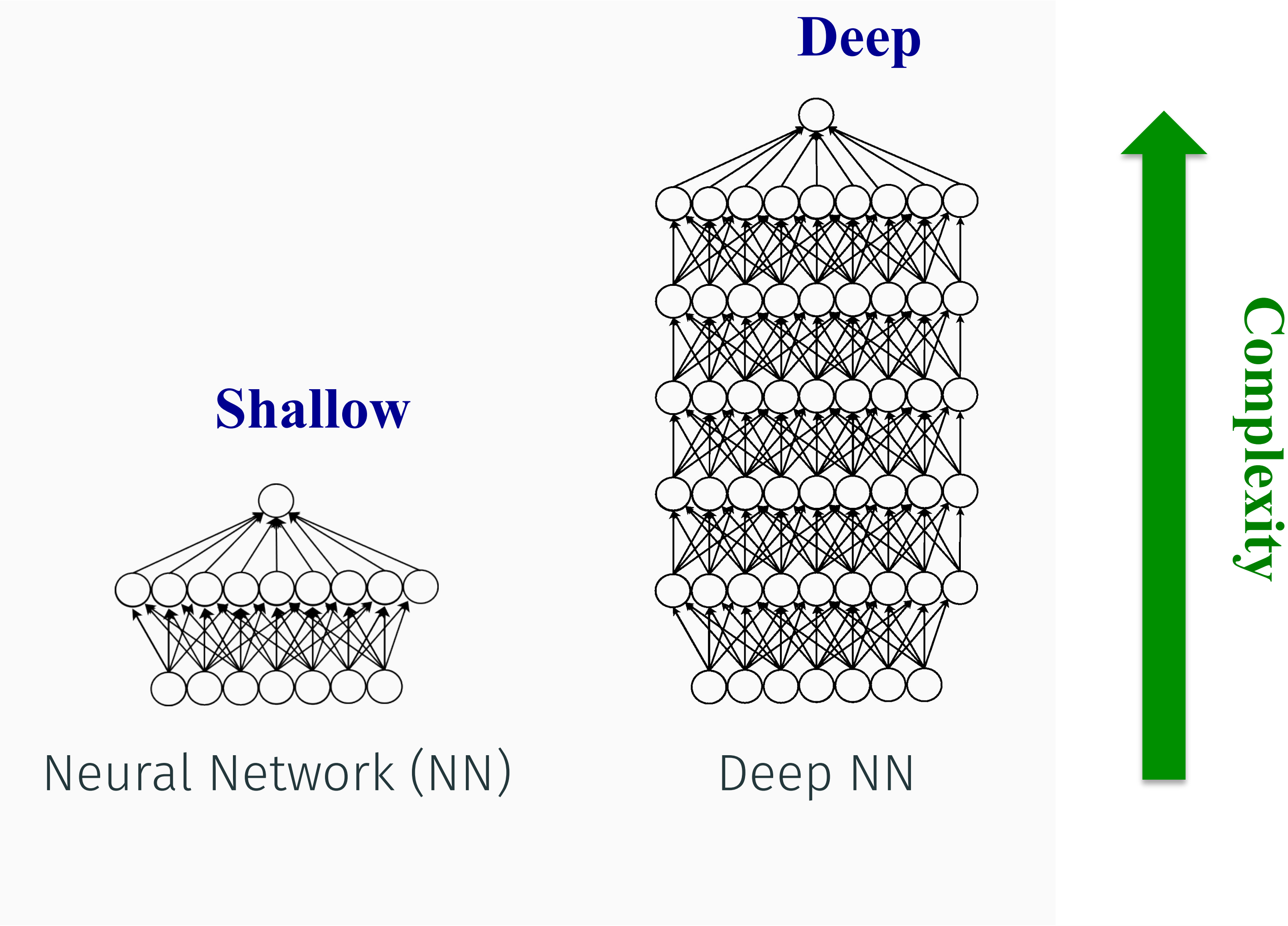
★ 3.5% 2016

# Deep Learning

## Deep Neural Networks (DNN) achieve significant performance improvements



# Deep Learning



# Deep Learning

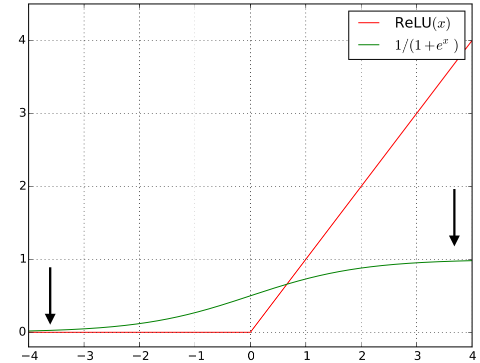
- **Training more complex models**
  - Increased Depth
  - Enlarged Width
  - Feedback/Convolution
  - Novel activation functions

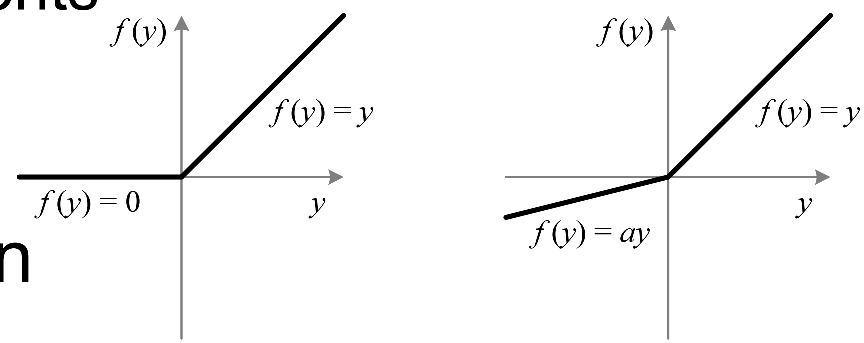
- **Effective strategies avoiding over-fitting**
  - Regularization

# ReLU

## Rectified Linear Unit (ReLU)

- Rectified neuron
- Faster training convergence
  - Better solutions than sigmoids
    - Vanishing gradients
  - Trained by back-propagation

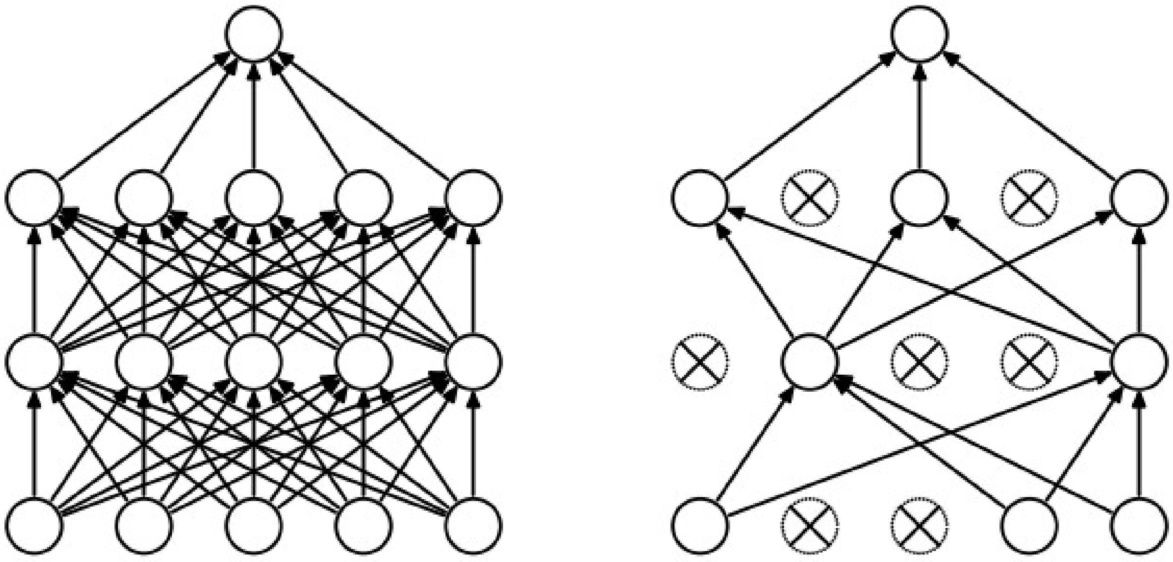



ReLU and Parametric PReLU

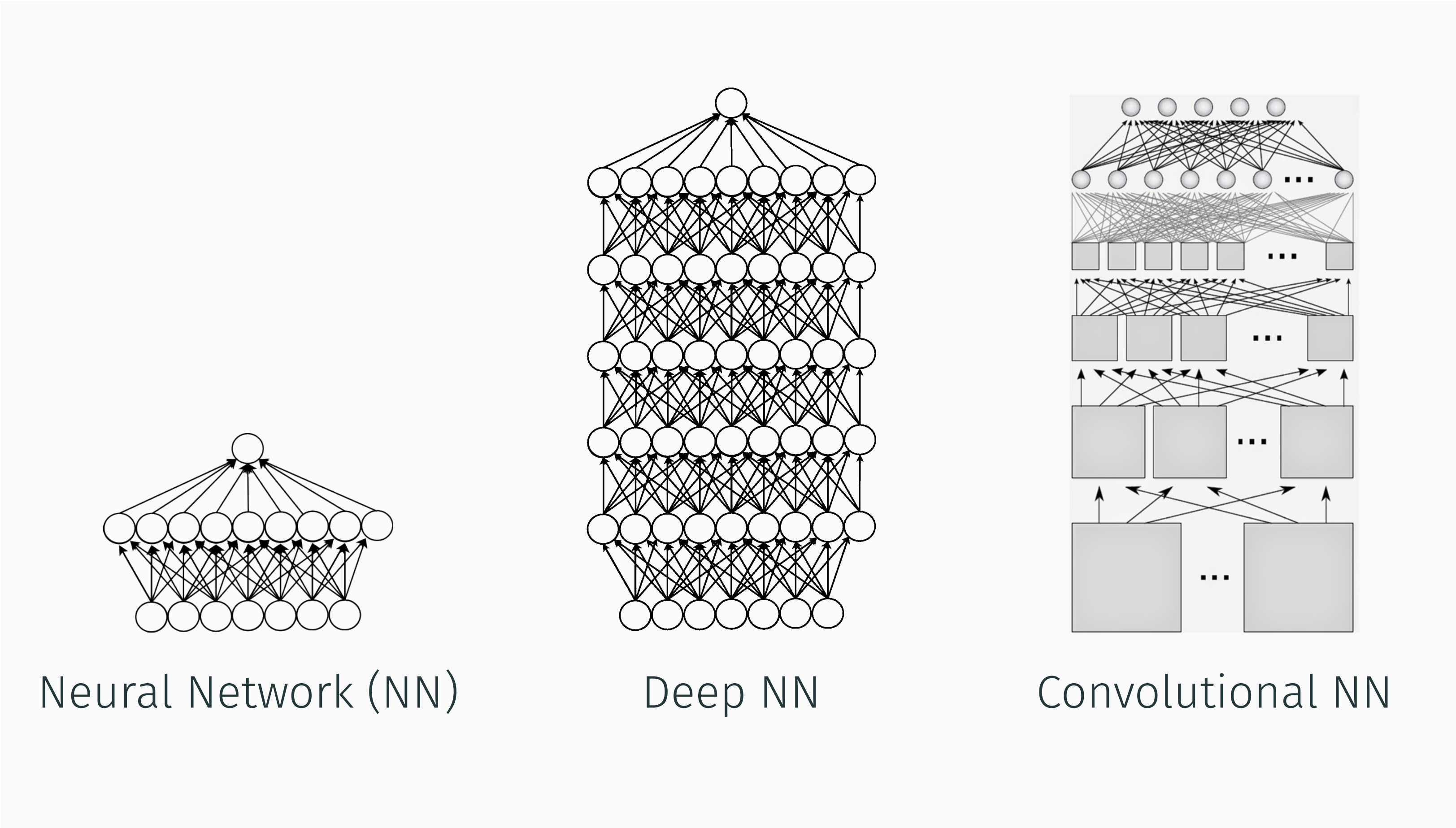
# Regularization

- i.e. Drop-Out



(a) Standard Neural Net (b) After applying dropout.

# Convolution

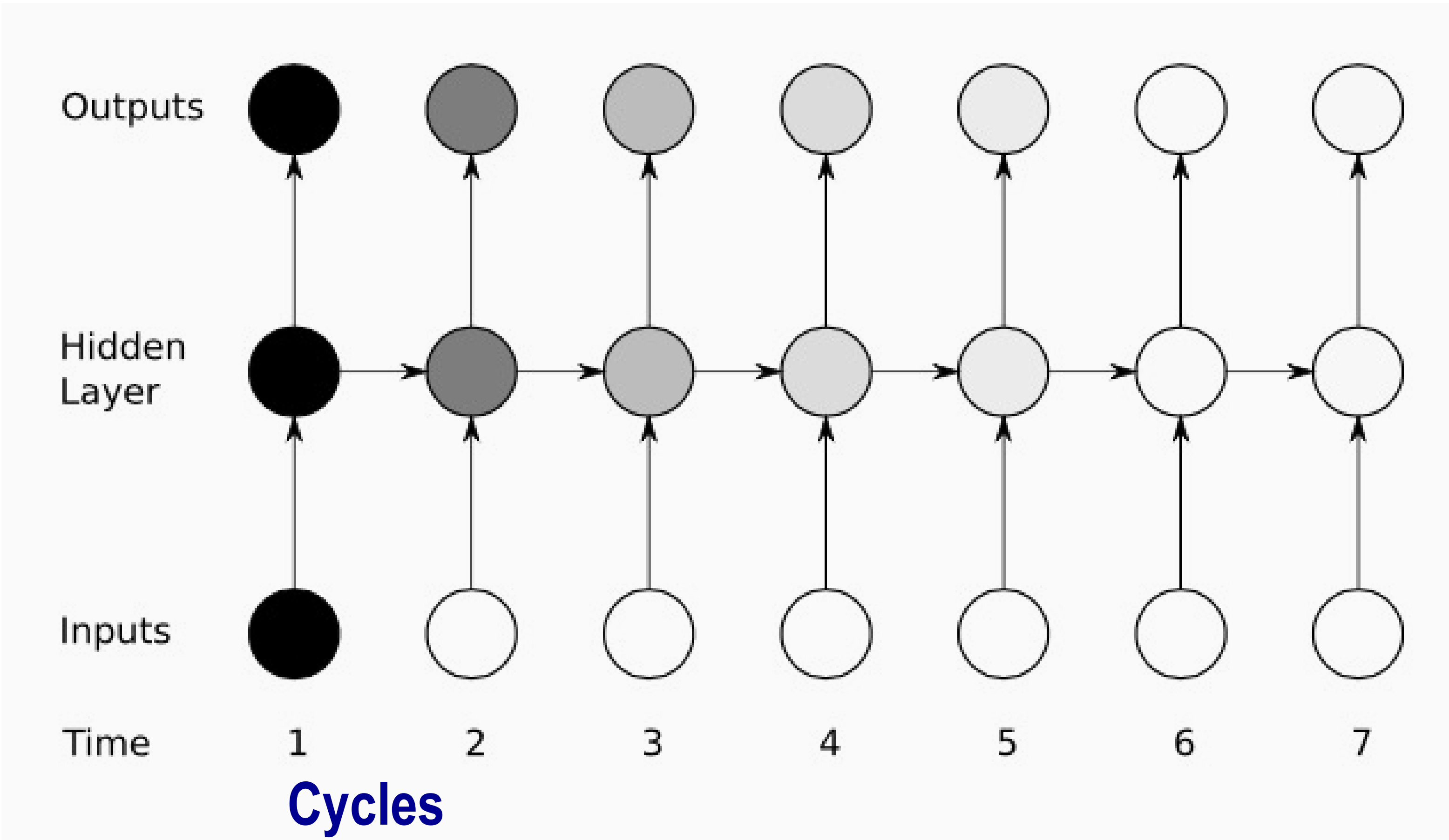


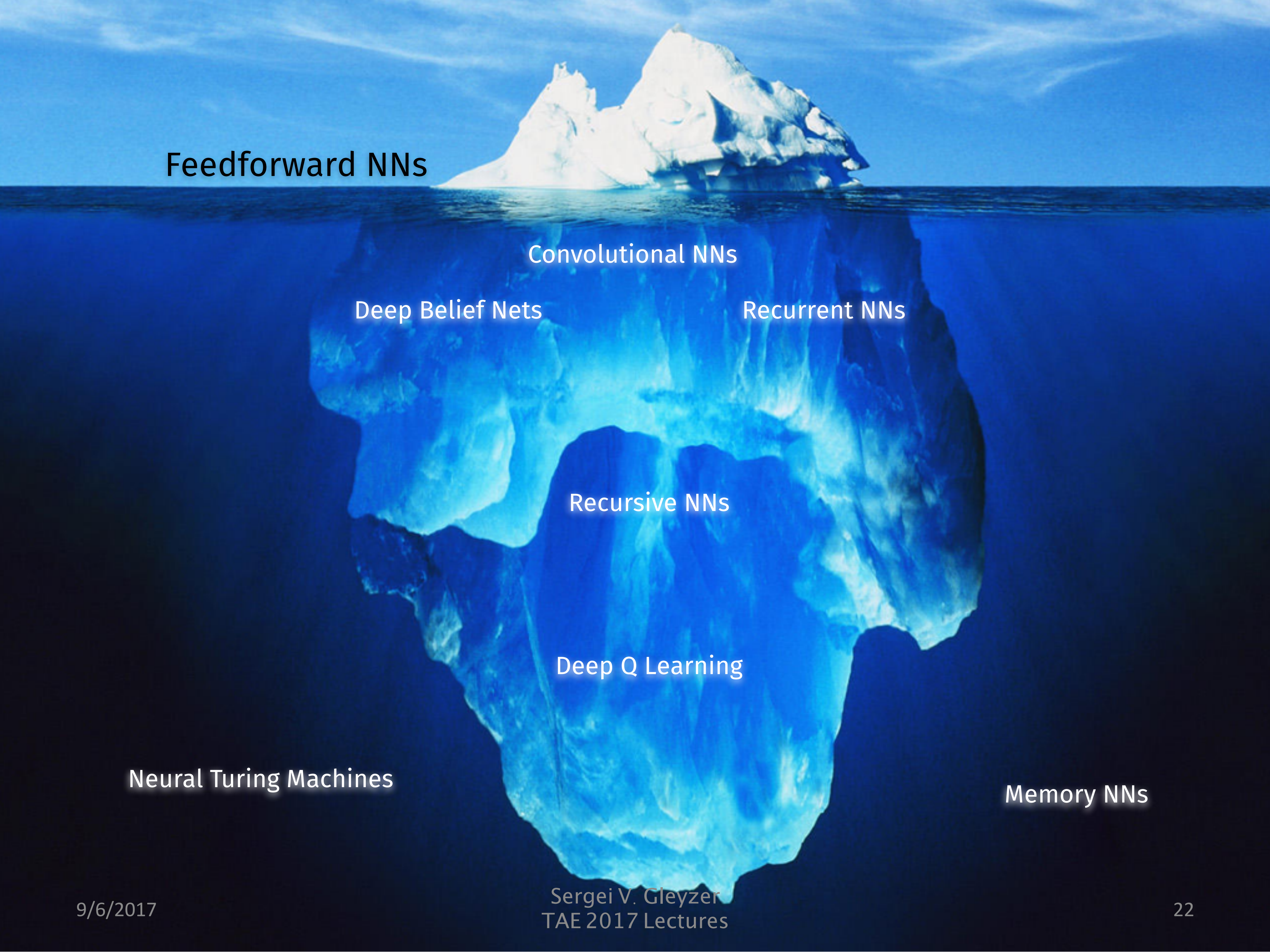
# Filters

## Convolutional Neural Networks:

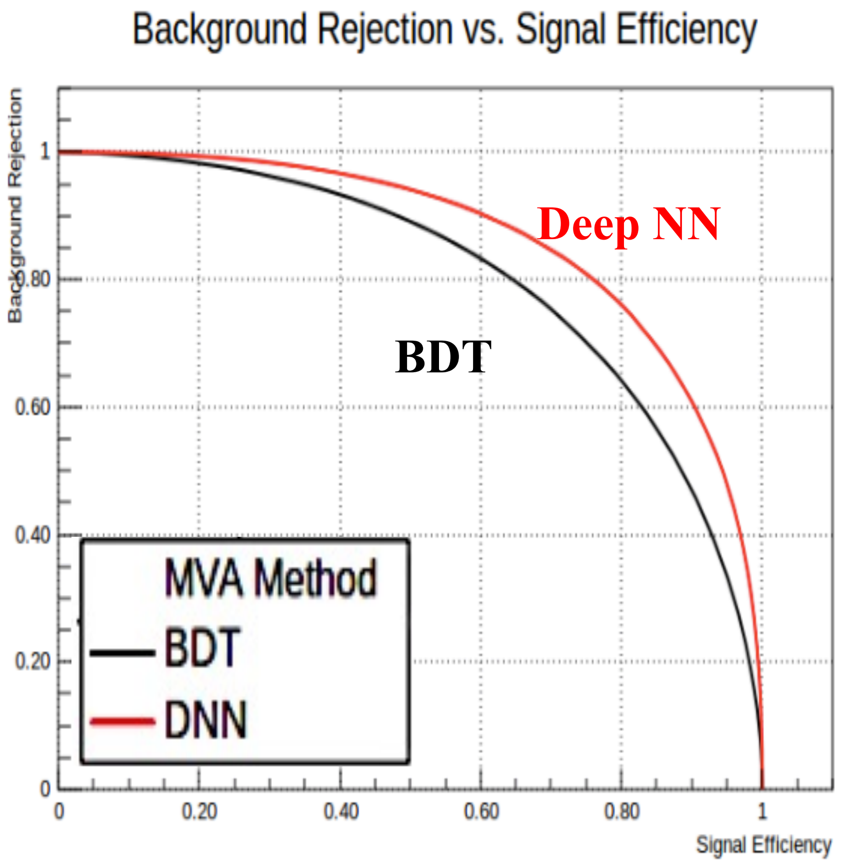
Unsupervised Feature Learning

# Recurrent NN



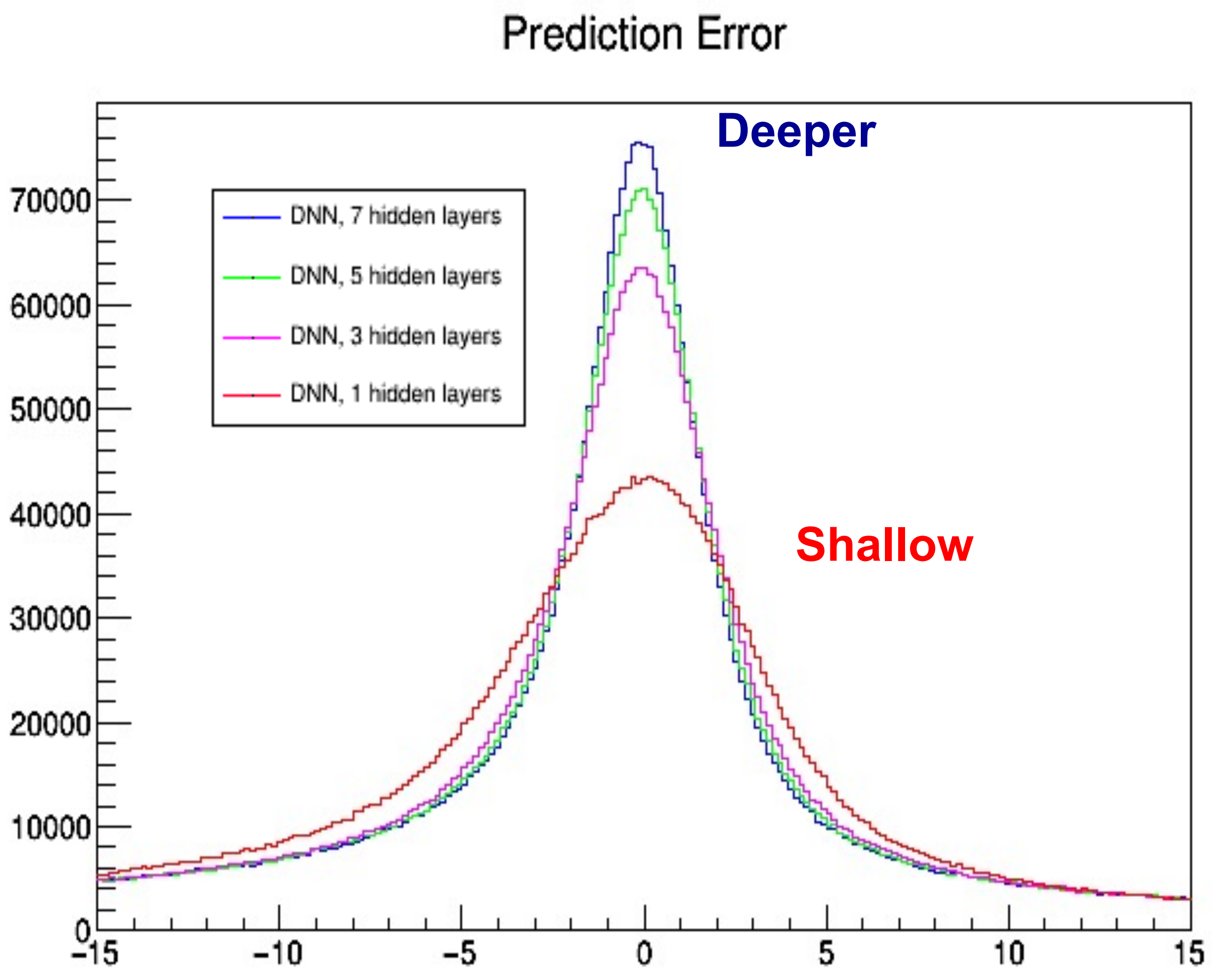

Feedforward NNs
Convolutional NNs
Deep Belief Nets
Recurrent NNs
Recursive NNs
Deep Q Learning
Neural Turing Machines
Memory NNs

# Deep Learning

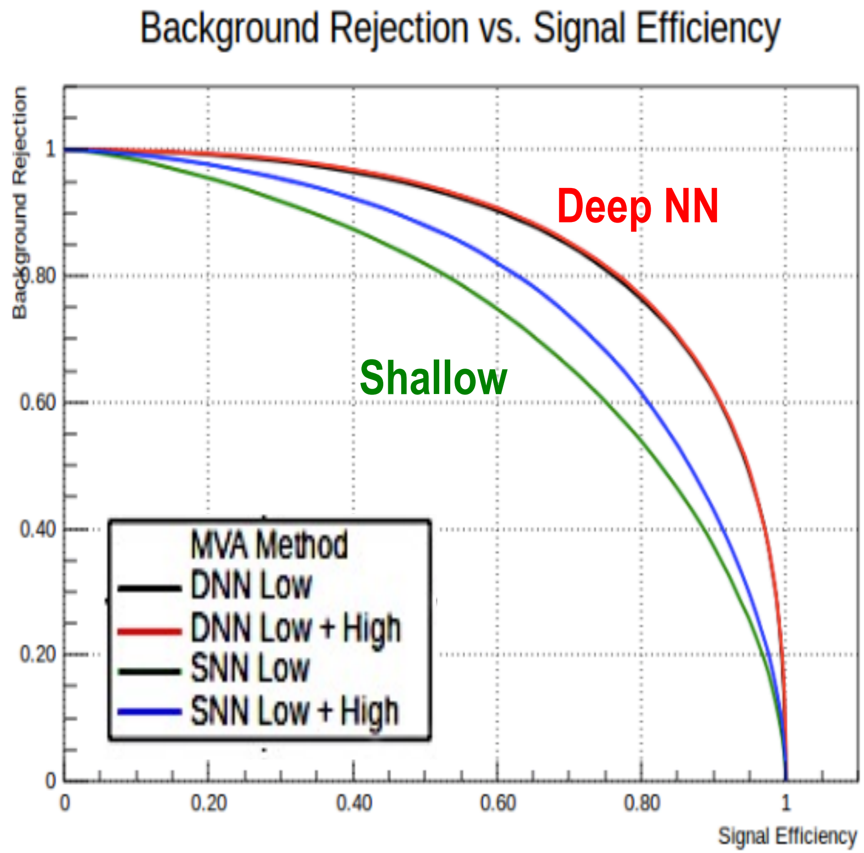



**Higher performance compared to previous ML methods**

# Deep Learning Regression



# Deep Learning




**Significant performance improvement in deep vs. shallow**

# Summary

**Machine learning is a powerful branch of data science**

- Many methods and applications
- Lectures covered basics and decision-tree based methods