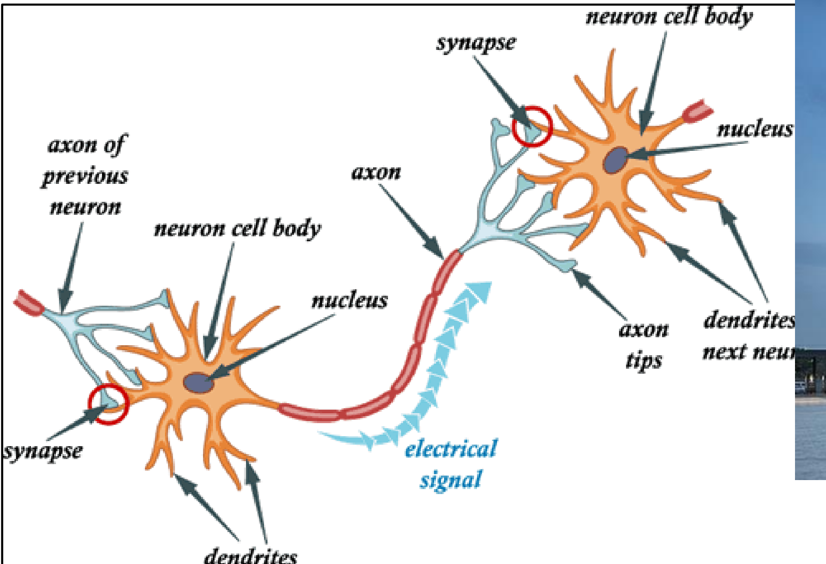
Deep Learning

(01076533)

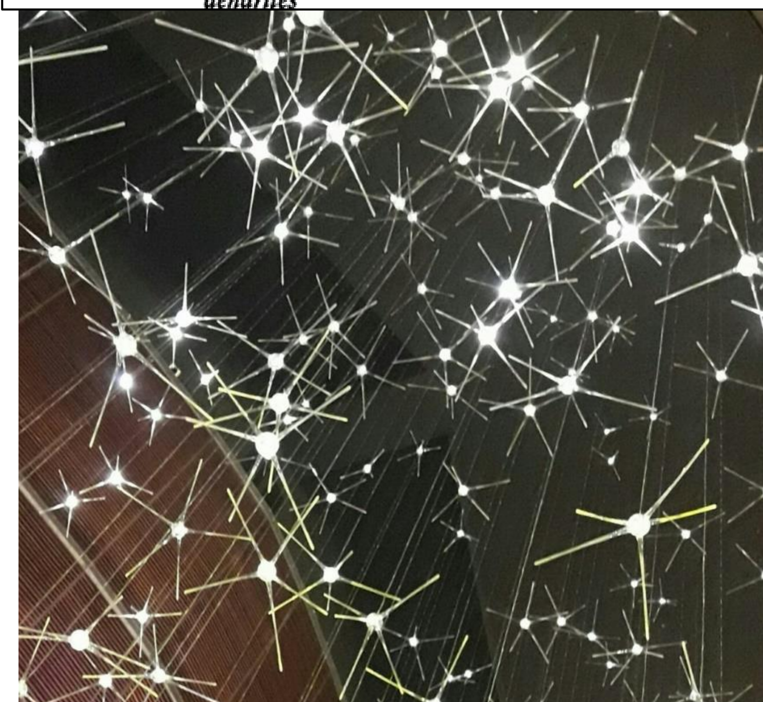
Kietikul Jearanaitanakij
Department of Computer Engineering, KMITL

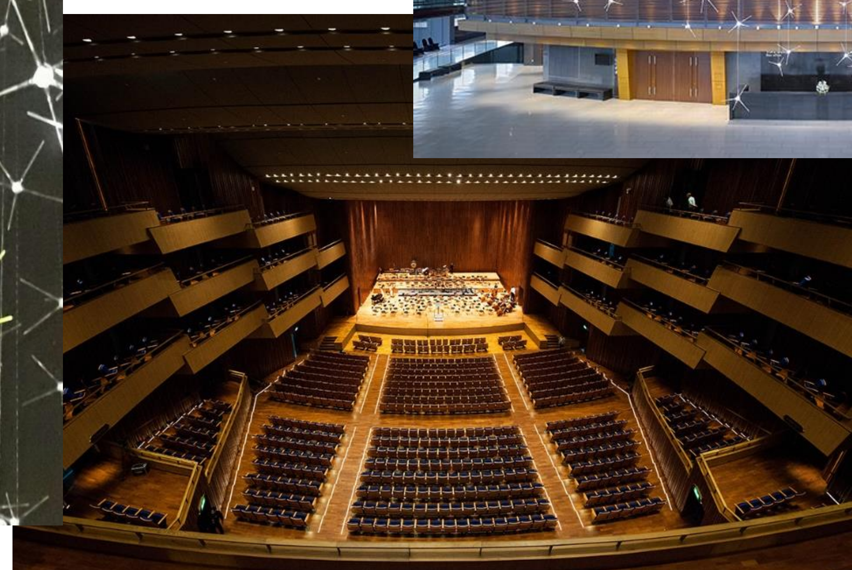
Some slides are adapted from

- cs231n: Convolutional Neural Networks for Visual Recognition @Stanford University


neuron cell body
synapse
nucleus
axon of
previous
neuron
axon
neuron cell body
nucleus
axon
tips
dendrites
synapse
electrical
signal
dendrites


มหิดลสิทธาคาร
PRINCE MAHIDOL HALL





Pre-requisites

- Python programming language + Numpy
- Linear algebra : N-dimensional matrix, matrix operations (add, subtract, multiplication, transpose, etc.)
- Calculus : Mostly derivatives.
- Optimization with gradient descent.

Course Outline

- Types of learning & Basic concepts
- Linear classification with perceptron
- Basic optimization with gradient descent
- Single-layer/Multilayer neural networks
- Keras part I : Single-layer & Multilayer NN + Assignment 1
- Loss function
- Backpropagation algorithm
- Convolutional neural network
- Simple case study of Convolutional neural network
- Keras Part II : Convolutional NN + Assignment 2
- Literature reviews of Convolutional NN

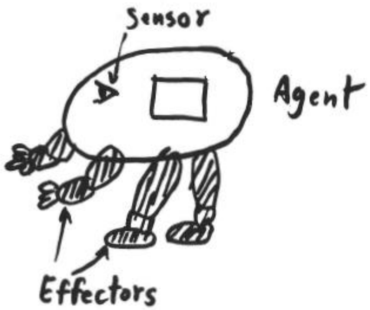
Grading

• Programming assignment 1	25%
• Programming assignment 2	25%
• Final exam	50%

Remarks:

- Programming assignments will be evaluated by walk-through grading.
- 4 students per 1 group.
- Python + Numpy + TensorFlow coding.

Types of learning



1. Unsupervised learning

– The agent learns patterns in the input even though no explicit label is supplied.

– The most common unsupervised learning task is clustering.

– For example, a taxi agent might gradually develop a concept of "Good traffic days" and "Bad traffic days" without ever being given labeled examples of each by a teacher.

  - X1: Rain, BigEvent, WorkDay, MonthEnd
  - X2: NonRain, NonBigEvent, NonWorkDay, NonMonthEnd
  - X3: Rain, NonBigEvent, WorkDay, NonMonthEnd
  - X4: NonRain, BigEvent, NonWorkDay, MonthEnd

A taxi agent may cluster X1, X4 as one cluster and X2, X3 as another cluster.

2. Reinforcement learning

- The agent learns from a series of reinforcements – rewards or punishments.
- For example, the lack of a tip at the end of the journey gives the taxi agent an indication that it did something wrong.
- It is up to the agent to decide which of the actions prior to the reinforcement were most responsible for it.

3. Supervised learning

- The agent observes some example input-output pairs and learns a function that maps from input to output.

- For example, from a taxi agent problem. By using given labeled “input-output” examples, a taxi agent will learn a function that maps form input to the outputs “Good traffic days” and “Bad traffic days”.

  - X1: Rain, BigEvent, WorkDay, MonthEnd, “Bad traffic days”
  - X2: NonRain, NonBigEvent, NonWorkDay, NonMonthEnd, “Good traffic days”
  - X3: Rain, NonBigEvent, WorkDay, NonMonthEnd, “Good traffic days”
  - X4: NonRain, BigEvent, NonWorkDay, MonthEnd, “Bad traffic days”

Overfitting

- Overfitting occurs when a statistical learning model describes random errors or noises instead of the underlying relationship.
  - A learning model that has been **overfit** will generally have poor predictive performance, as it can exaggerate minor fluctuations in the data.
  - A good learning model should be **generalize** to both unseen and training data.

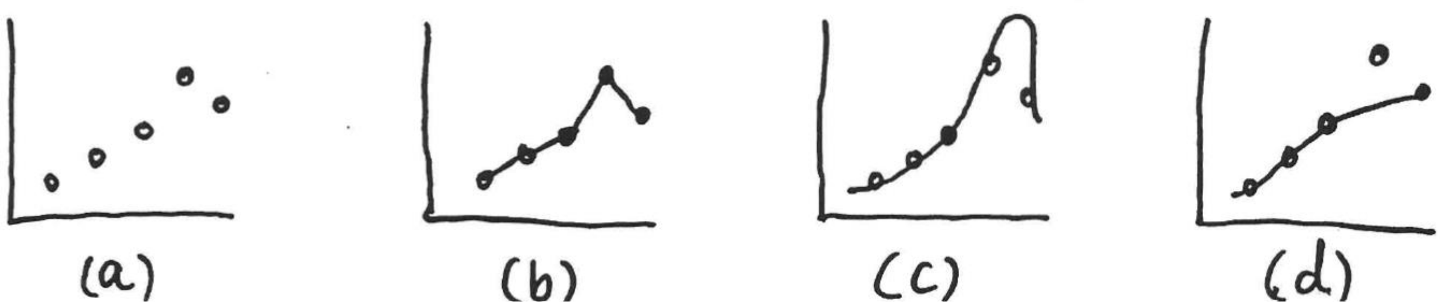


Fig. a : Training data
Fig. b and c tend to overfit the training data
Fig. d tends to generalize the training data (good hypothesis)

Evaluating the hypothesis

- We can simply use the error rate to check the quality of the hypothesis.

- $\text{Error rate} = \dfrac{\text{Number of examples which are wrong predicted}}{\text{Total number of examples}}$

- However, a hypothesis h has a low error rate on the training set does not mean that it will generalize well.

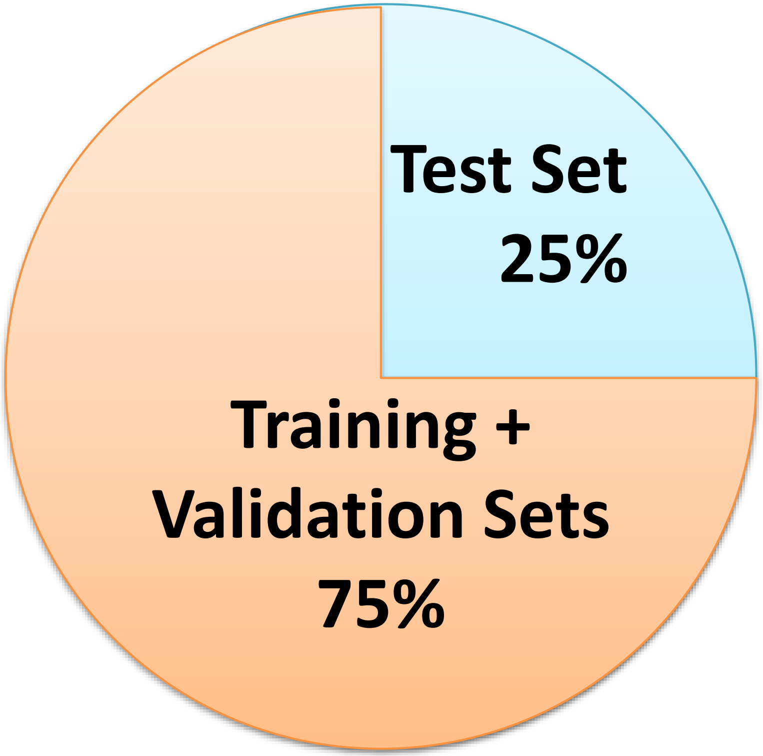
- To get the more accurate evaluation, we need to test the hypothesis on an unseen set (test set) of examples.

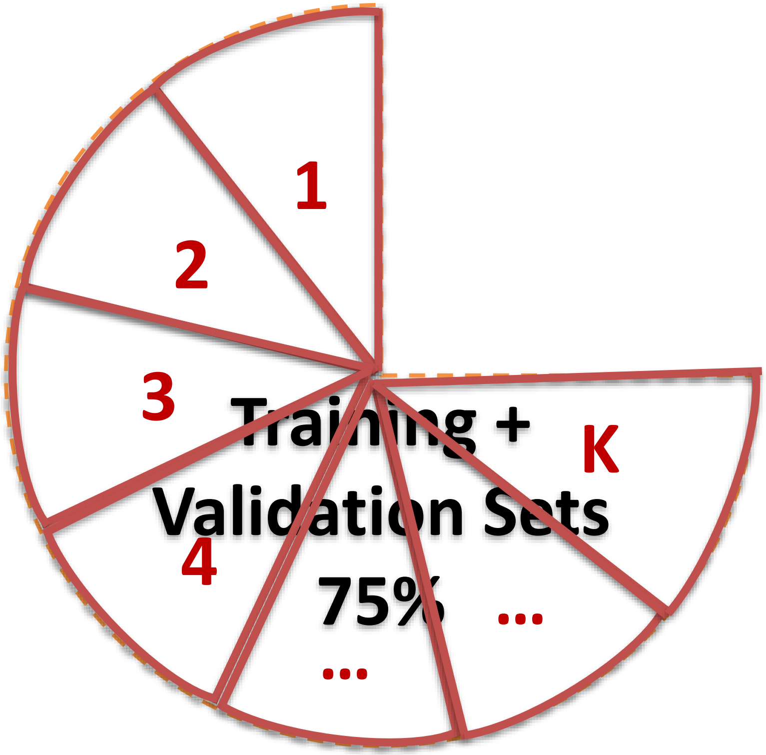
- **Holdout cross-validation** :

Randomly split the available data into a training set (e.g. 75%) and a test set (e.g. 25%).

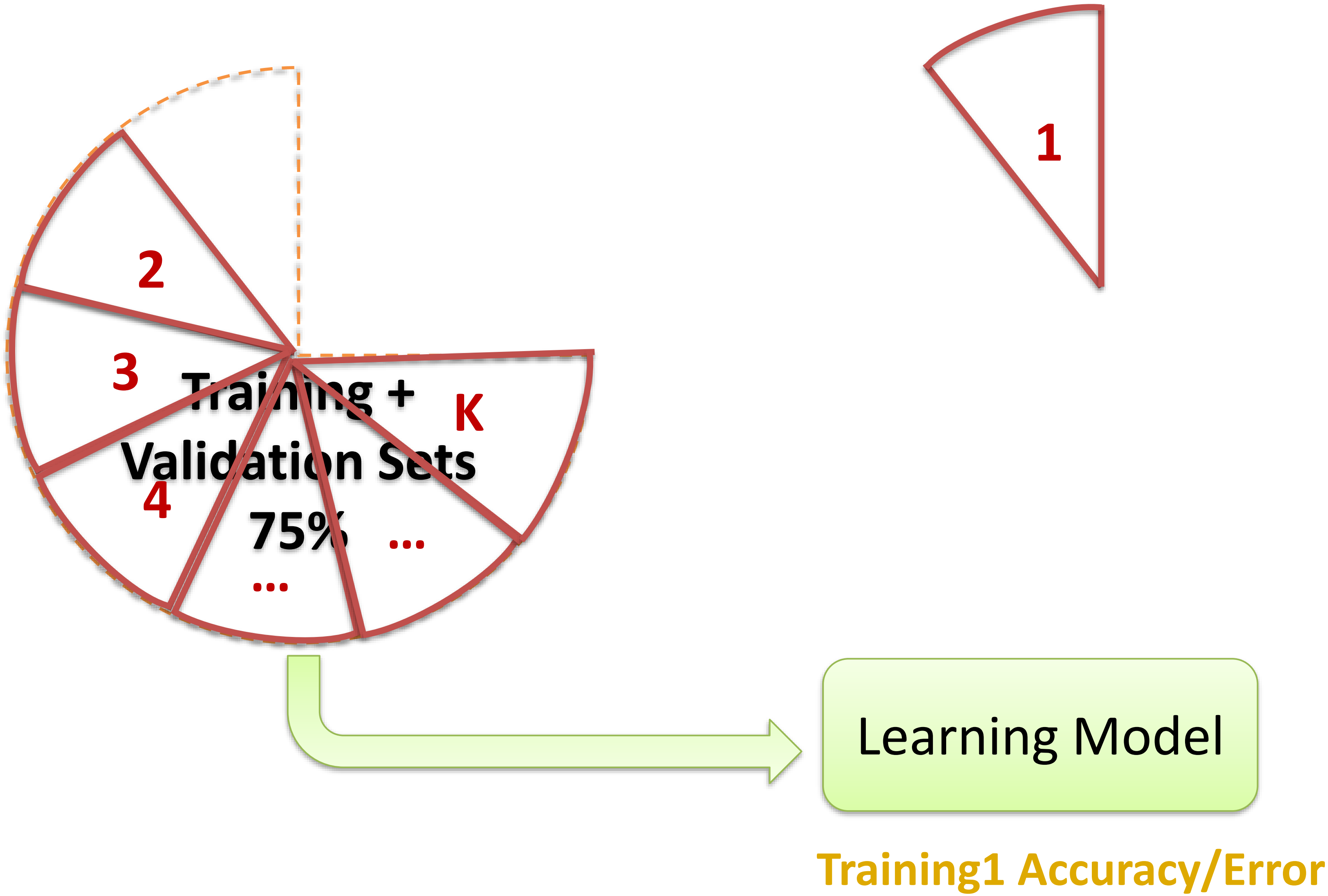
- **K-fold cross-validation**:
  - Split the training data into k equal subsets.
  - Then perform k rounds of learning. On each round, 1/k of the data is held out as a validation set and the remaining examples are used as training data.
  - The average error rate of the k rounds is finally calculated. Popular values for k are 5 and 10. The extreme is k=n (the total number of examples), also known as leave-one-out cross-validation or LOOCV.

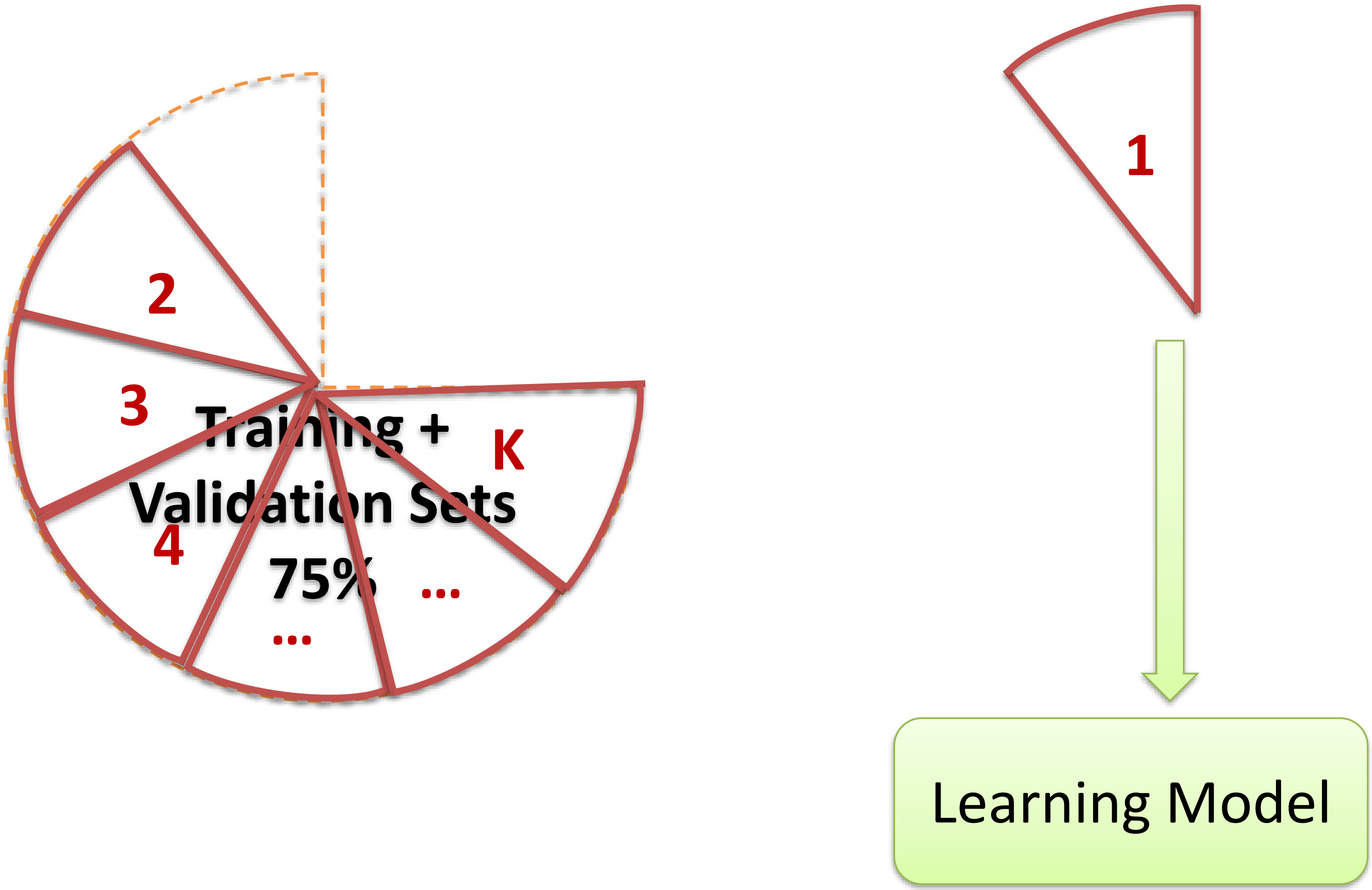
K-Fold Cross Validation





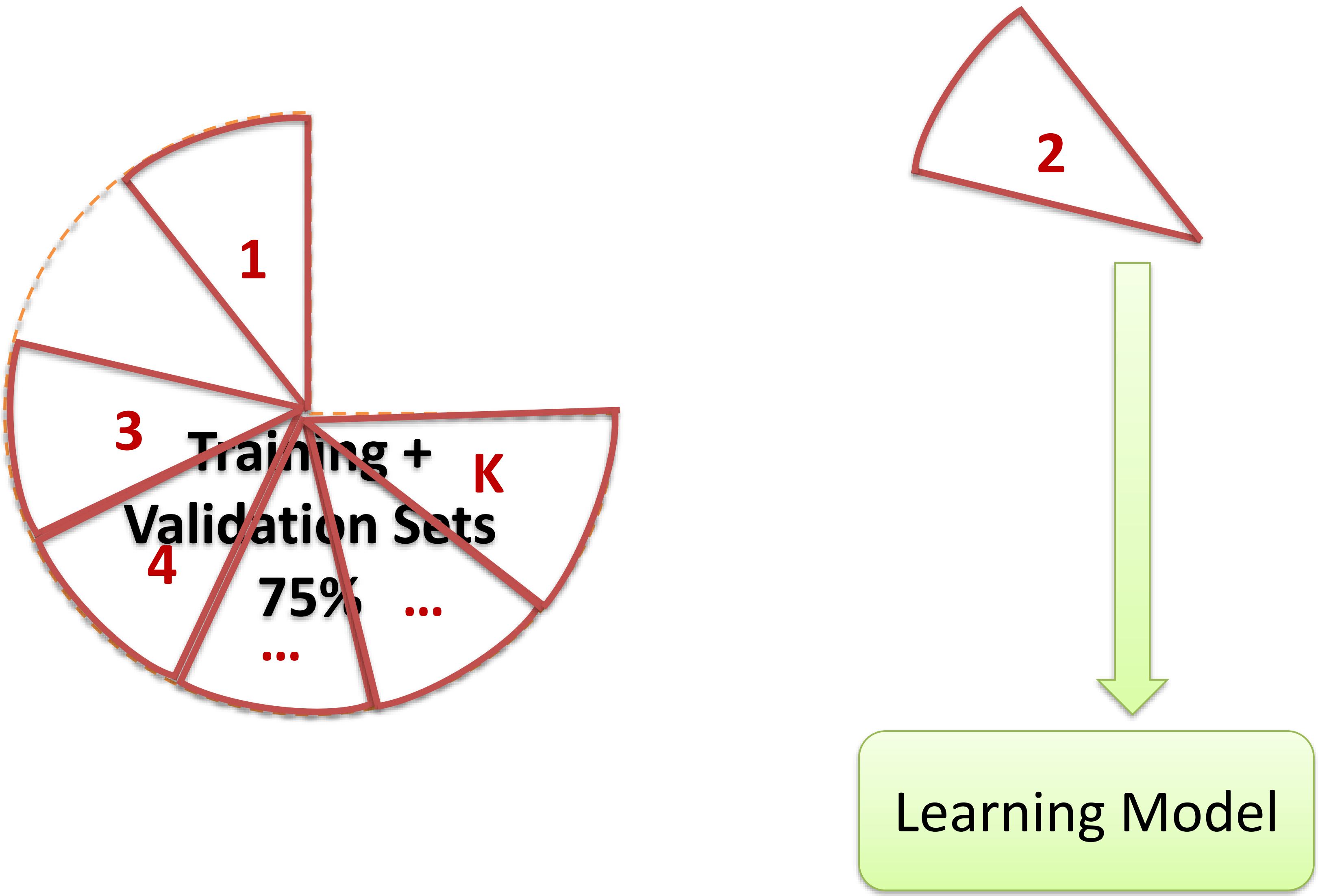
Learning Model


1
2
3
Training +
Validation Sets
75%
K
4
...
...
Learning Model
Training1 Accuracy/Error

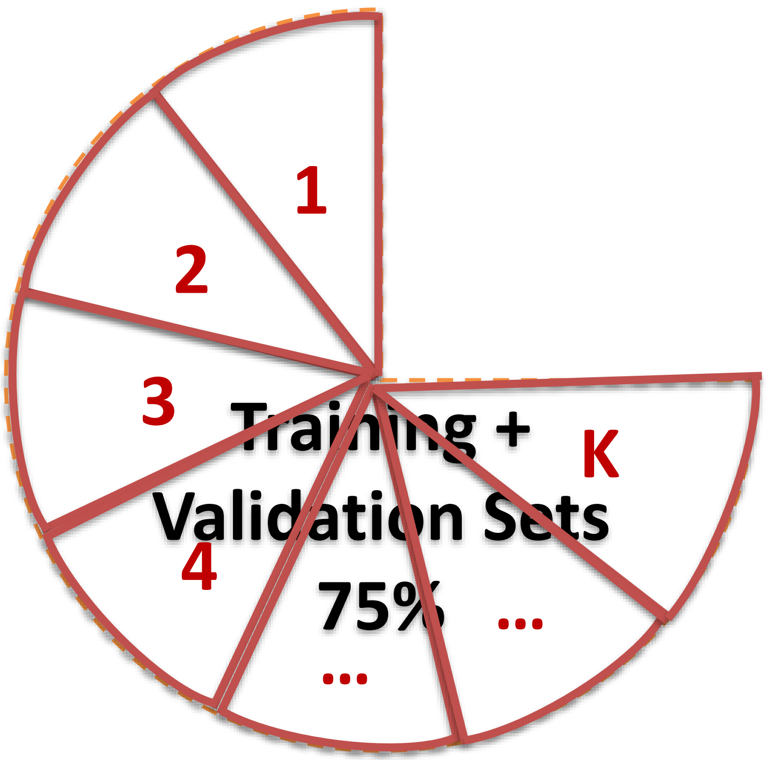


Validation1 Accuracy/Error

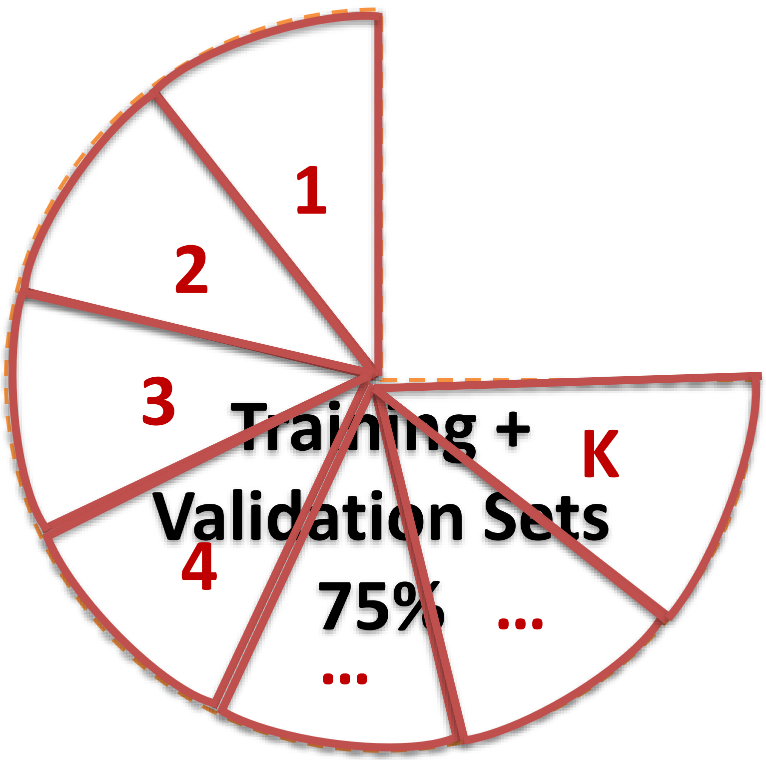
Training2 Accuracy/Error



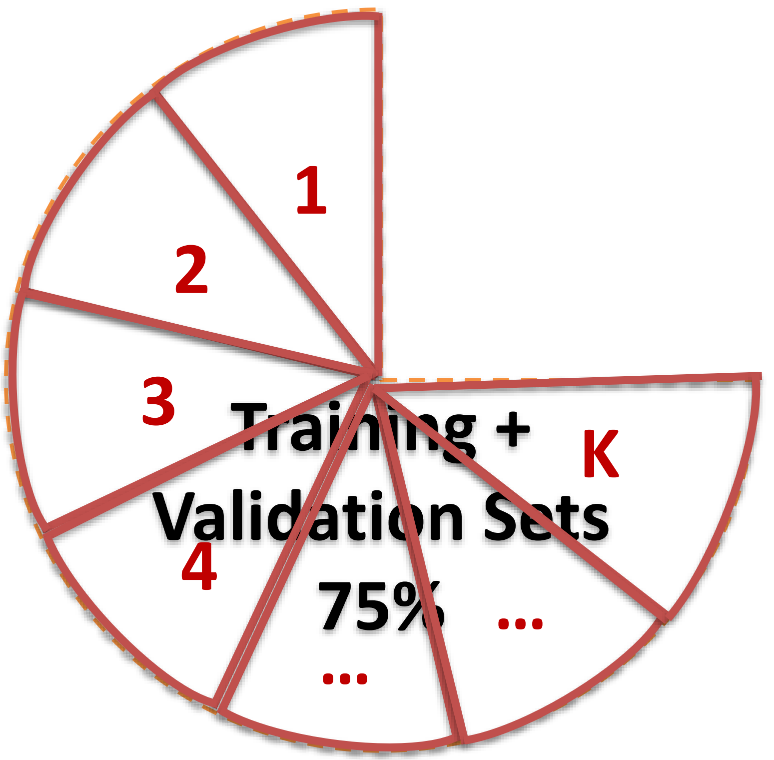
Validation2 Accuracy/Error



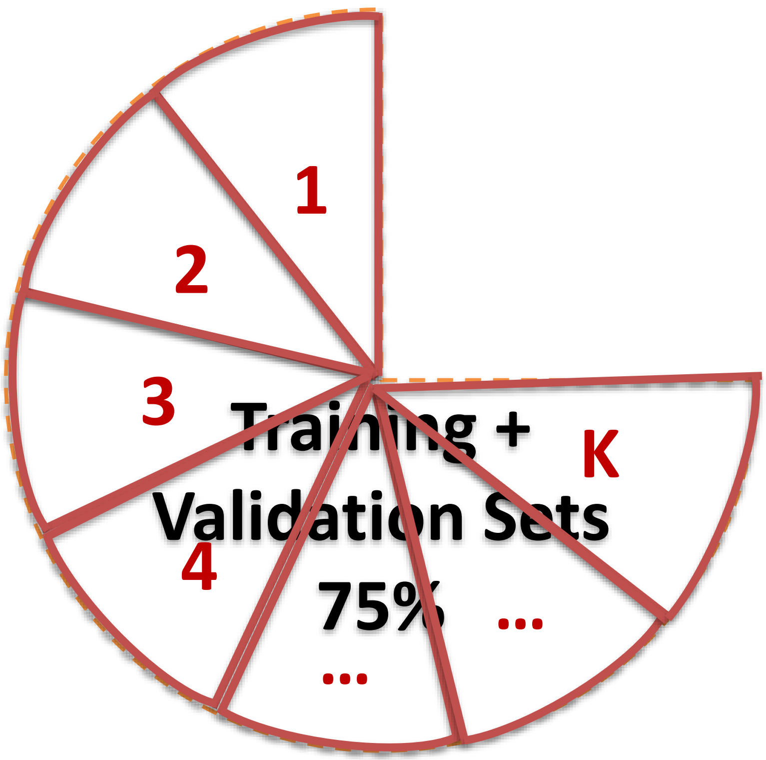
Learning Model



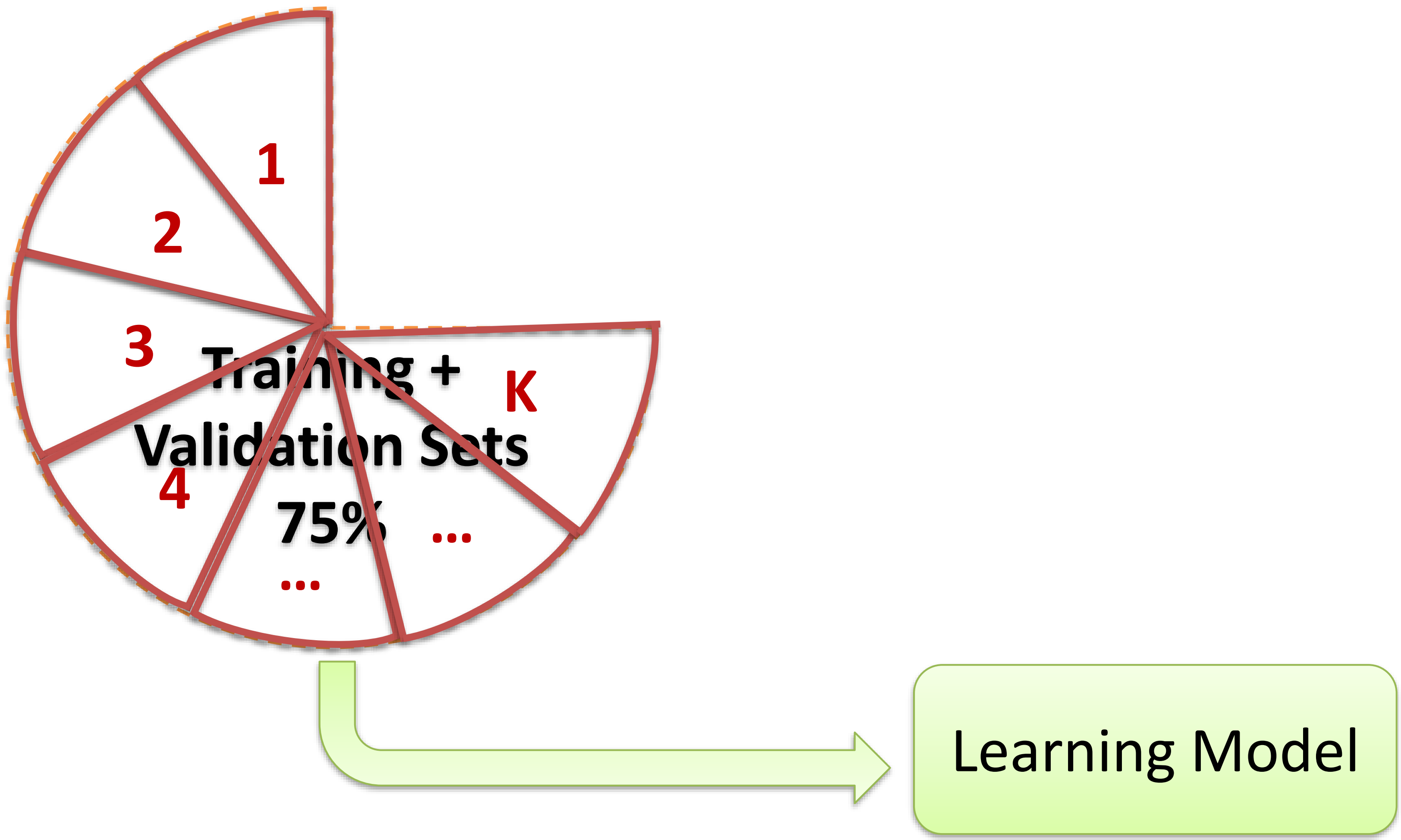
Learning Model



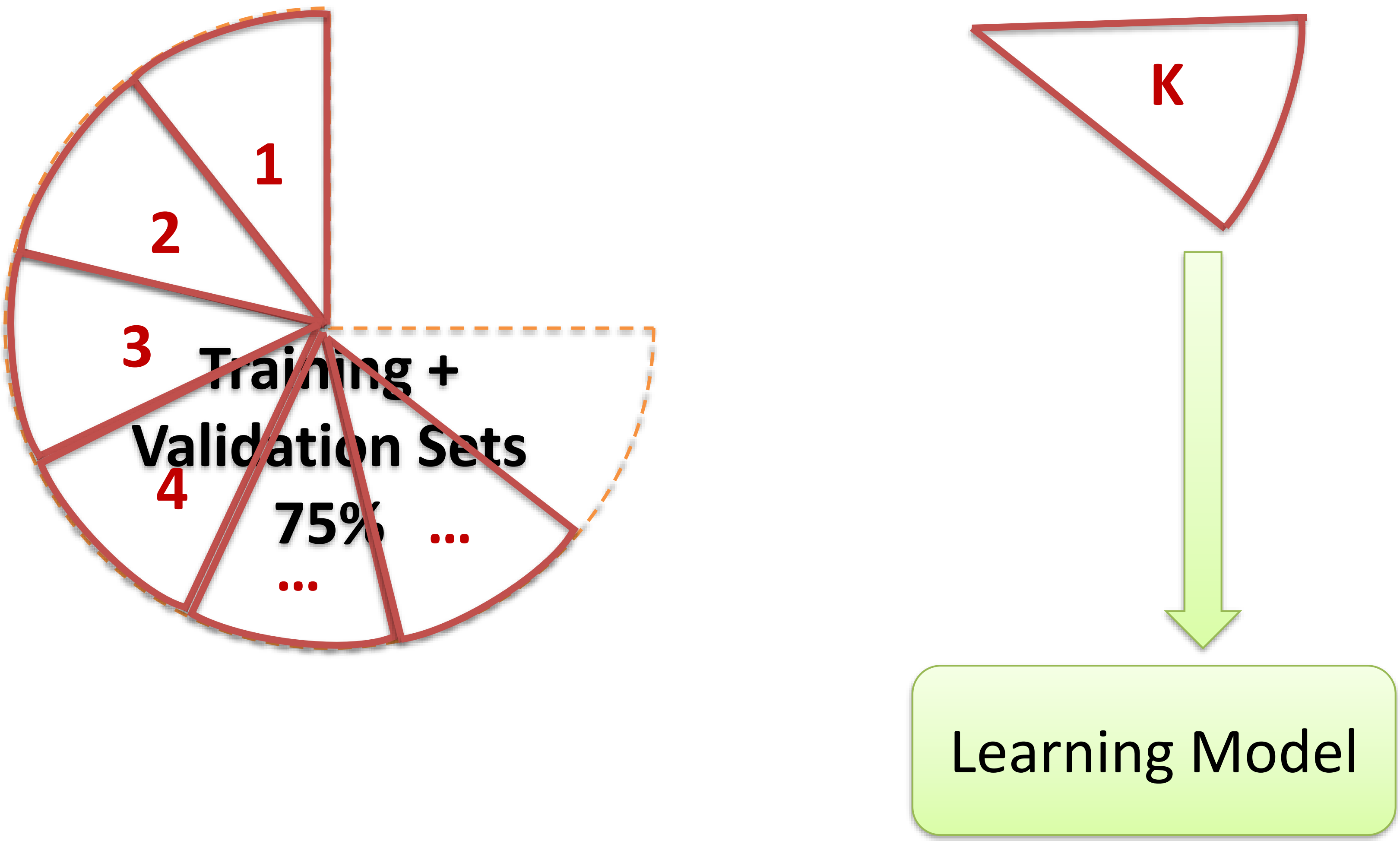
Learning Model



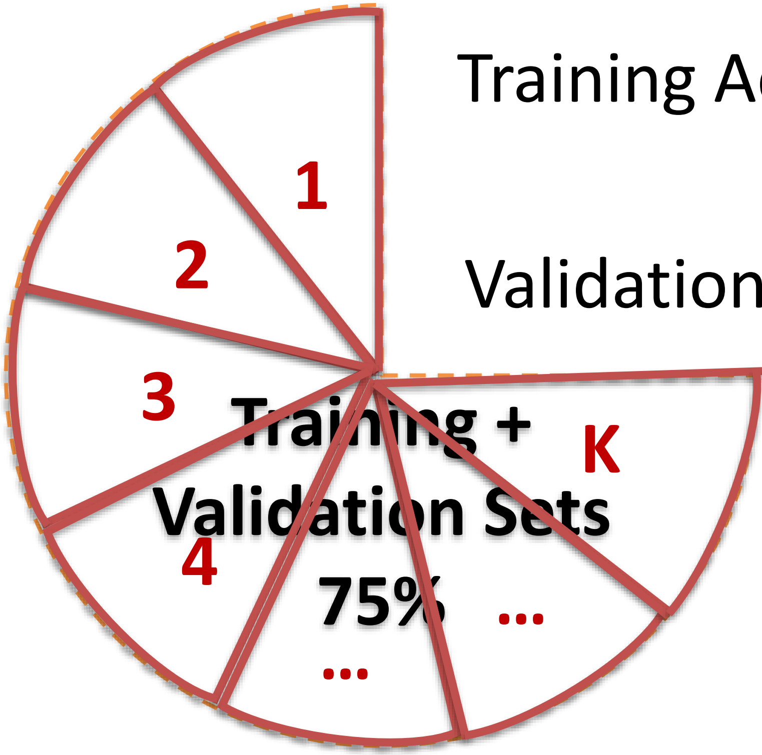
Learning Model



Training K Accuracy/Error

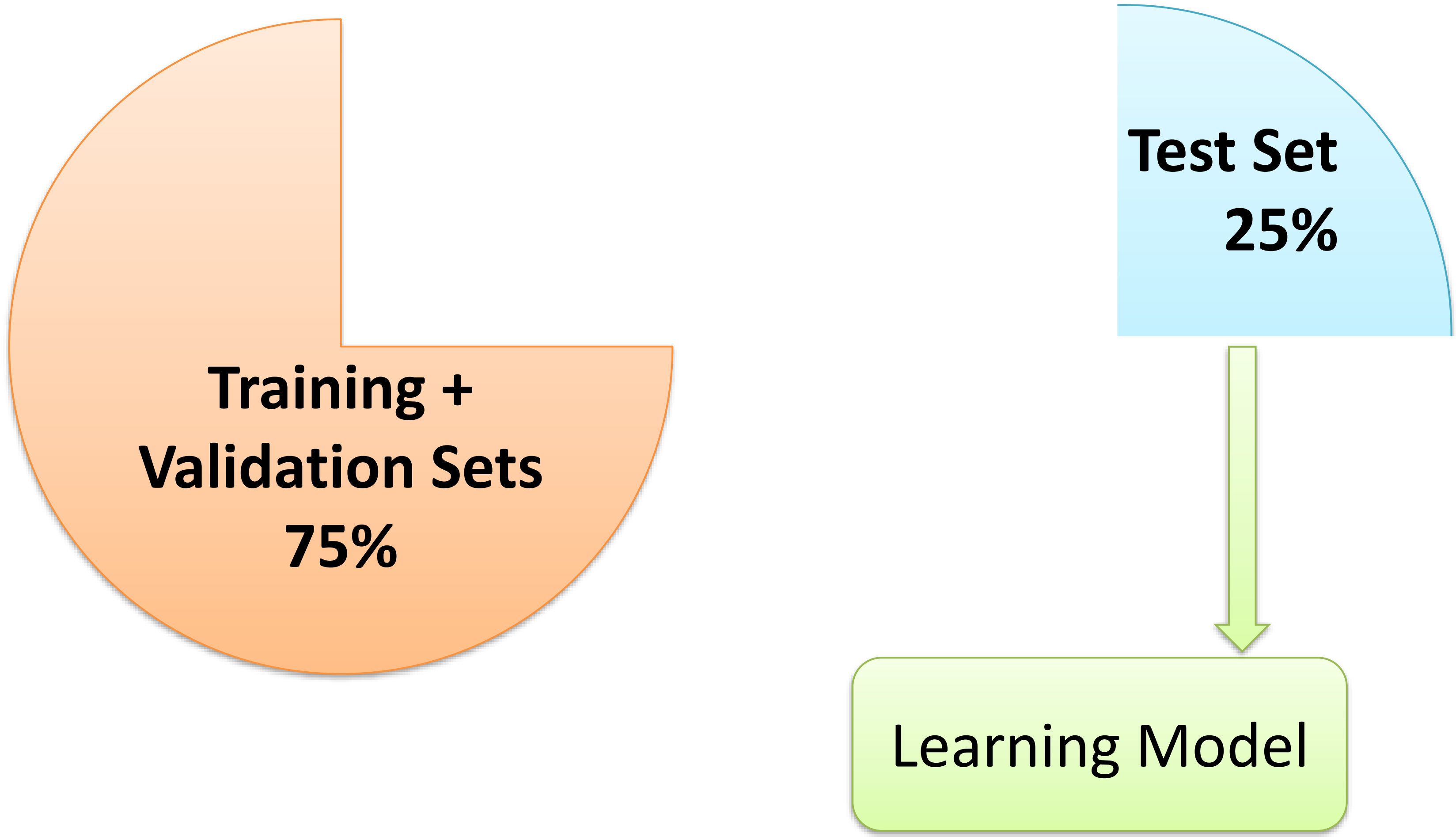


Validation K Accuracy/Error



Training Accuracy = $(\sum_{i=1}^{K} Train\_i\_Acc)$/K

Validation Accuracy = $(\sum_{i=1}^{K} Valid\_i\_Acc)$/K



Test Accuracy/Error

Note: Percentage of test set can be varied, depends on how many total of instances we have.

Deep learning
(from wikipedia)

Deep learning (also known as **deep structured learning** or **hierarchical learning**)

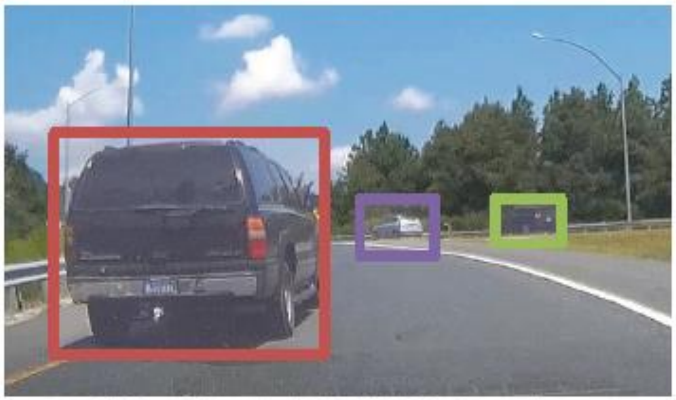
- Use artificial neural networks (ANNs) to learning tasks that contain more than one hidden layer.
- Deep learning is part of a broader family of machine learning methods based on learning data.
- Learning can be supervised, partially supervised or unsupervised.

Most popular applications of deep learning is computer vision.

Other applications are speech recognition, natural language processing, audio recognition, social network filtering, machine translation and bioinformatics.

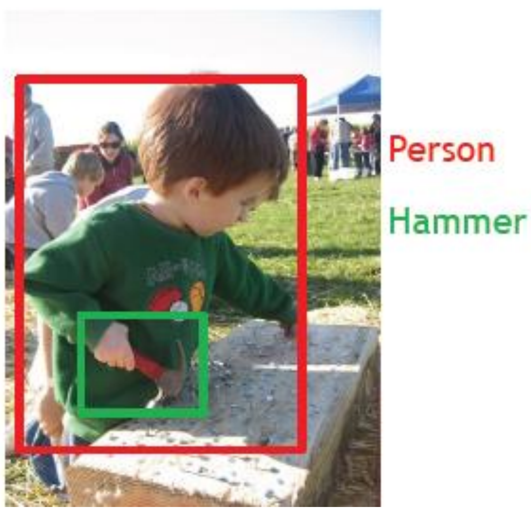
This course will mainly focus on the convolutional neural network for the classification problem on image datasets.

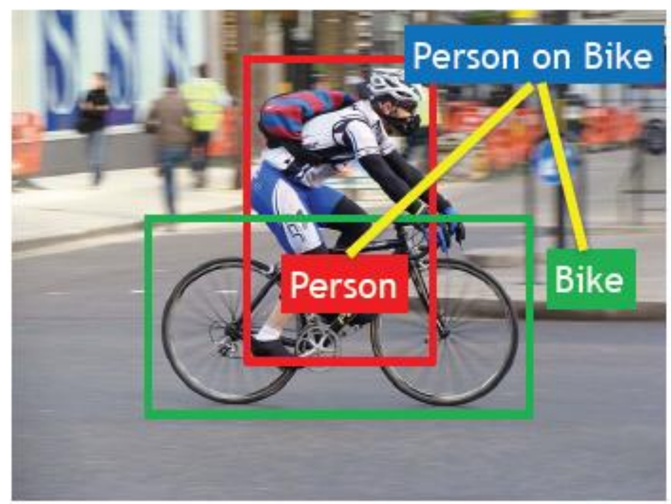
There is a number of visual recognition problems that are related to image classification, such as object detection, image captioning.



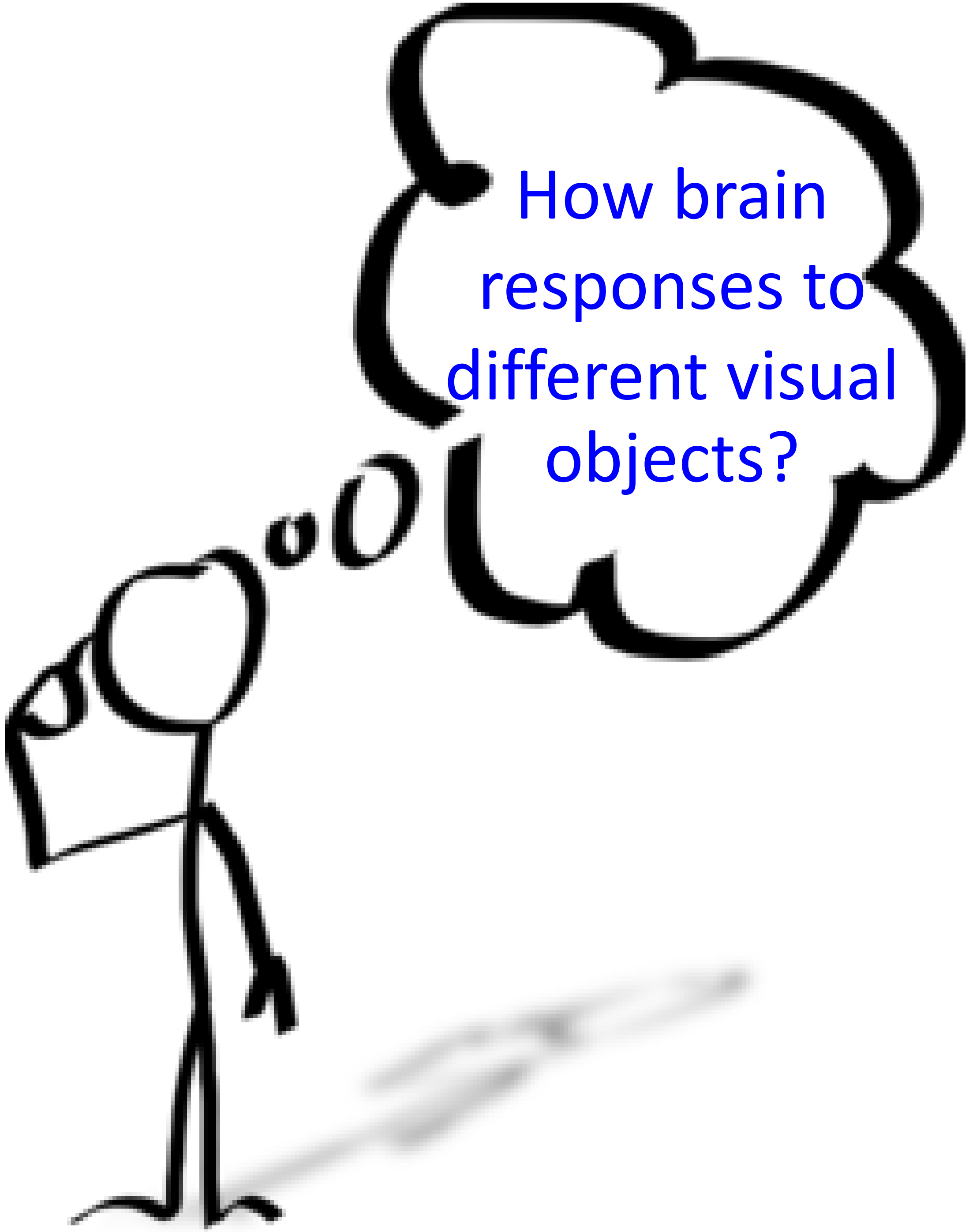
https://youtu.be/yQwfDxBMtXg

- Object detection
- Action classification
- Image captioning
- ...

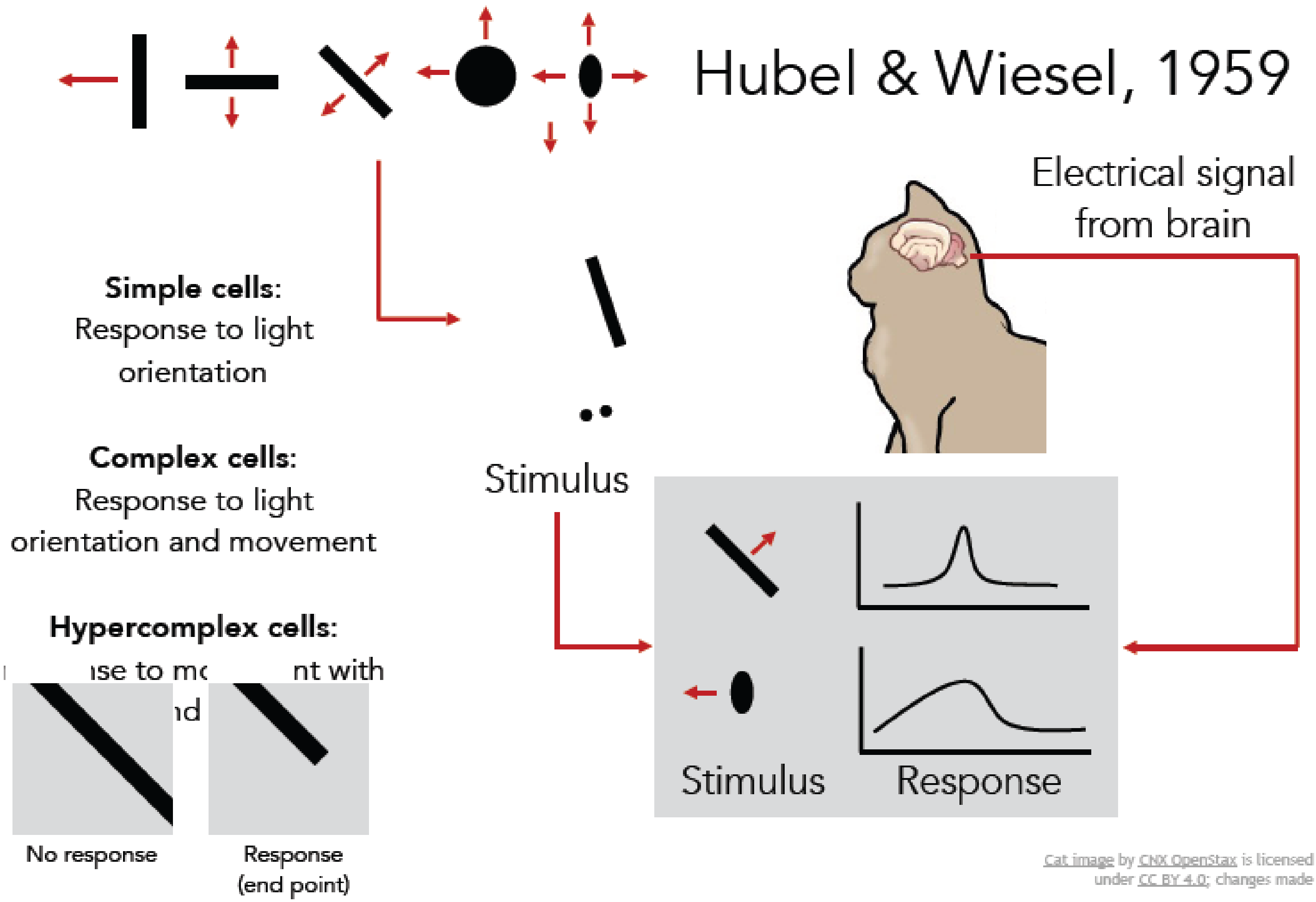




Images from cs231n @Stanford


How brain responses to different visual objects?

Visual Cortex Study

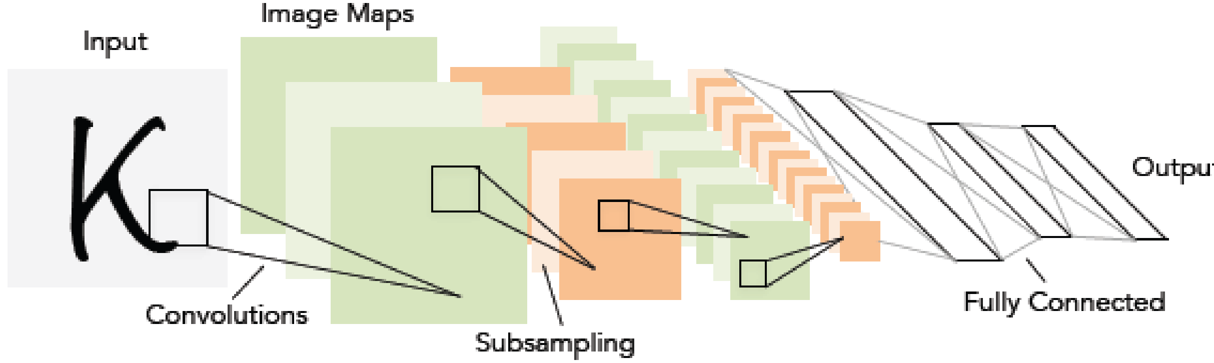


https://youtu.be/IOHayh06LJ4 https://youtu.be/jw6nBWo21Zk

As you will see later in the class, convolutional neural network imitates visual cortex in the brain

Evolution of Convolutional Neural Networks (CNN)

1998
LeCun et al.



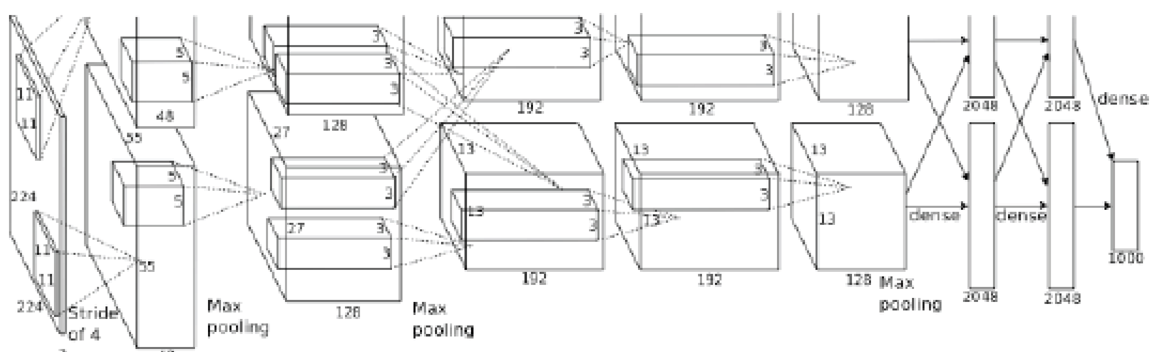
# of transistors

10^6

# of pixels used in training

10^7 NIST

2012
Krizhevsky et al.



# of transistors

10^9

GPUs

# of pixels used in training

10^{14} IMAGENET

- Next class
  - Python & Numpy for Deep Learning