

Deep Learning 101

How to train a neural network

Schedule

week	Date	Topic
9	10.27	Environment setup, python, Jupyter, PyCharm, TensorFlow, & regression
10	11.03	Training and testing
11	11.11	CNN
12	11.18	RNN
13	11.24	Autoencoder & GAN

Today's Class

- Recap
- Concepts:
 - Feature
 - Hypothesis
 - Cost function
 - Gradient descent
 - Backpropagation
 - Softmax
 - Cross entropy
 - One hot encoding
- To train a neural network
 - Dataset
- Lab time

Recap

- Neural network as a function
 - $y = f(x)$
- Perceptron
 - $Y = WX + b$
 - Two inputs: x_1, x_2
 - One output: y
 - Linear regression
- XOR problem
 - Linear regression can't solve the XOR problem
 - Require multivariate regression

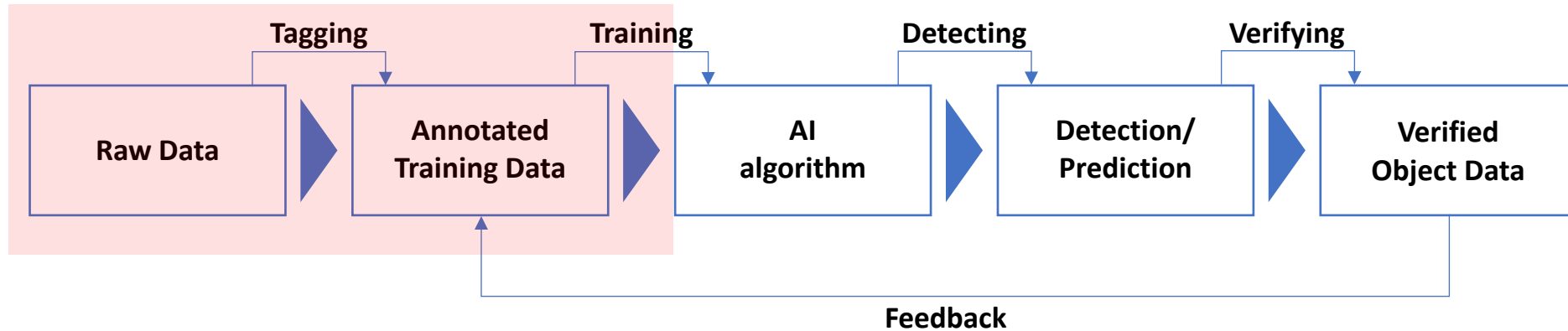
What

- Dataset: data for training and testing
 - Requires preprocessing
- Model: What the network learns
 - Training, validation, and testing
- Inference: Model in action
 - Predicting based on the learned model

The Challenge



The Time-Consuming Part for AI development



Training data is the most important part of AI development, but it is also the most difficult and time-consuming part

How to train a model

- Define input and output
- Decide on the input features
- Build layers of the network: hyperparameters
 - Number of layers
 - Learning rate
 - Number of epochs
 - Etc.
- Train the model: parameters
 - Weights and biases
 - Variables in TensorFlow
- Verify the model:
 - Using verification data

Models and Functions

- Hypothesis:

$$H(x_1, x_2, x_3) = w_1x_1 + w_2x_2 + w_3x_3 + b$$

- Activation:
 - Sigmoid, ReLU, LeakyReLU, etc.

- Cost:
$$cost(W, b) = \frac{1}{m} \sum_{I=1}^m (H(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}) - y^{(i)})^2$$

Matrix multiplication

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \times \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{bmatrix} = \begin{bmatrix} 58 & \\ & \end{bmatrix}$$

The "Dot Product" is where we **multiply matching members**, then sum up:

$$(1, 2, 3) \bullet (7, 9, 11) = 1 \times 7 + 2 \times 9 + 3 \times 11 \\ = 58$$

<https://www.mathsisfun.com/algebra/matrix-multiplying.html>

Functions using matrix

- Hypothesis

- $Y = WX + b$
$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = (x_1 w_1 + x_2 w_2 + x_3 w_3)$$

- Activation function

- Cost function

$$cost(W, b) = \frac{1}{m} \sum_{I=1}^m (H(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}) - y^{(i)})^2$$

Activation functions

- Introduces non-linearity
- Normalizes the output: activation functions are also called Normalization functions
- Different kinds
 - Step function: $f(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{w} \cdot \mathbf{x} + b > 0, \\ 0 & \text{otherwise} \end{cases}$
 - Sigmoid:

$$S(x) = \frac{1}{1 + e^{-x}}$$

$S(x)$ = sigmoid function

e = Euler's number

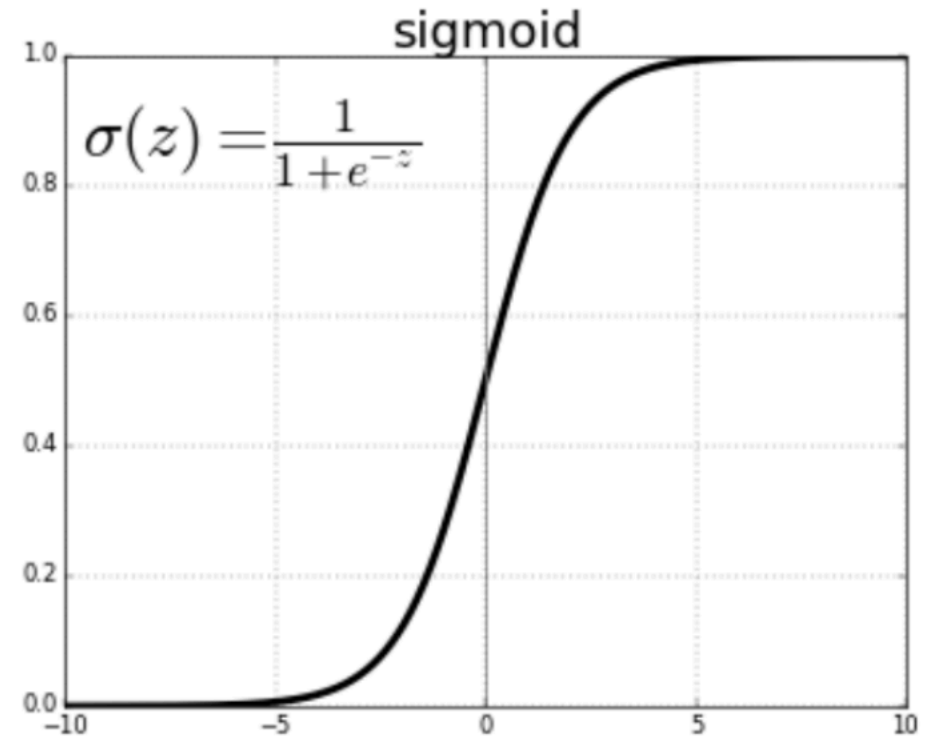
Activation function: Sigmoid

- Sigmoid:

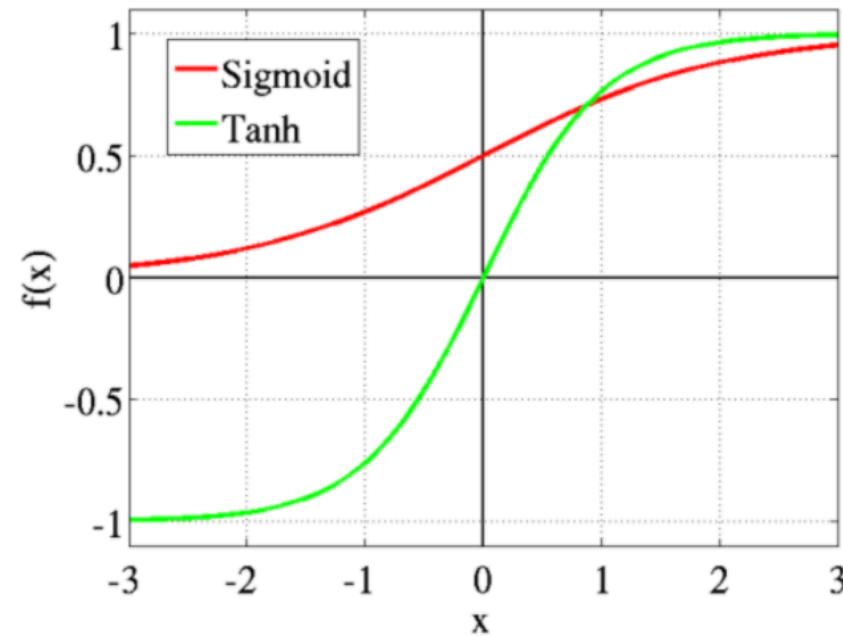
$$S(x) = \frac{1}{1 + e^{-x}}$$

$S(x)$ = sigmoid function

e = Euler's number



Activation functions: Sigmoid and Tanh



<https://www.neuronactivator.com/blog/what-even-is-activation-function>

Activation Functions: ReLU and Leaky ReLU

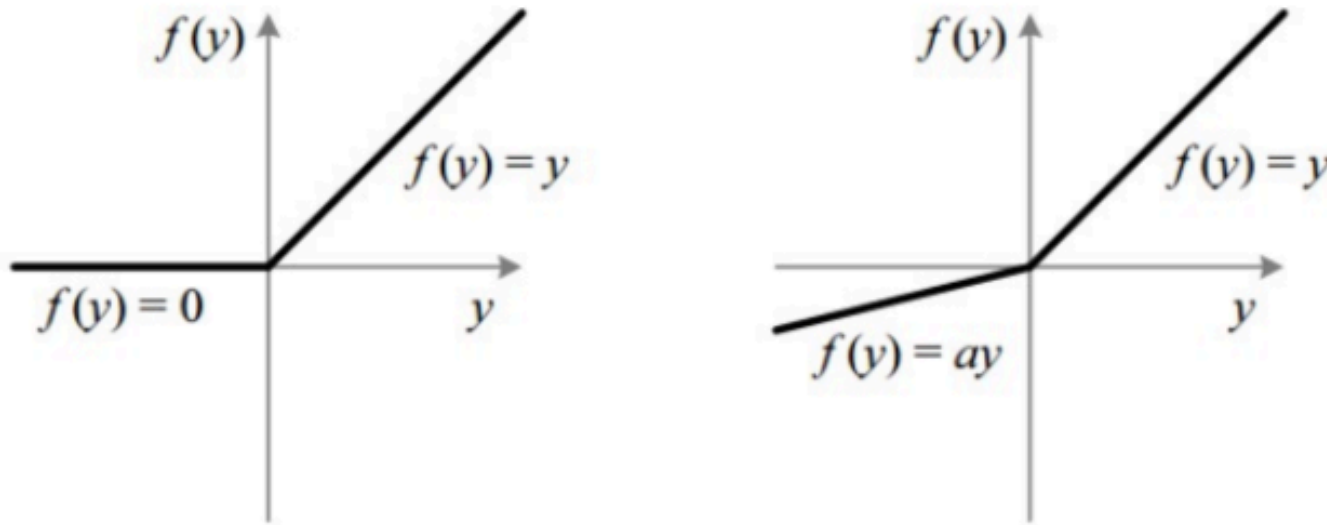


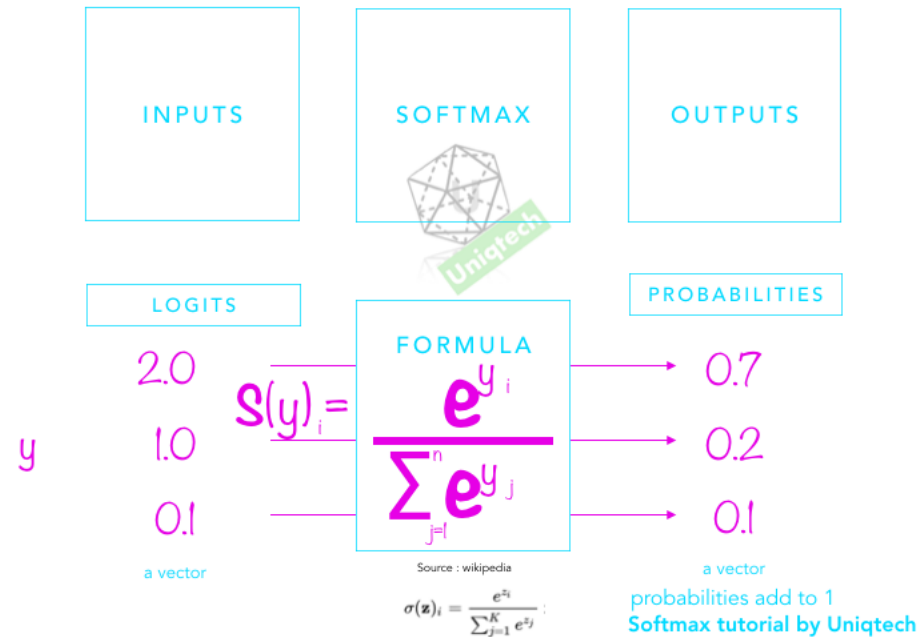
Fig : ReLU v/s Leaky ReLU

<https://www.neuronactivator.com/blog/what-even-is-activation-function>

Activation Function: Softmax

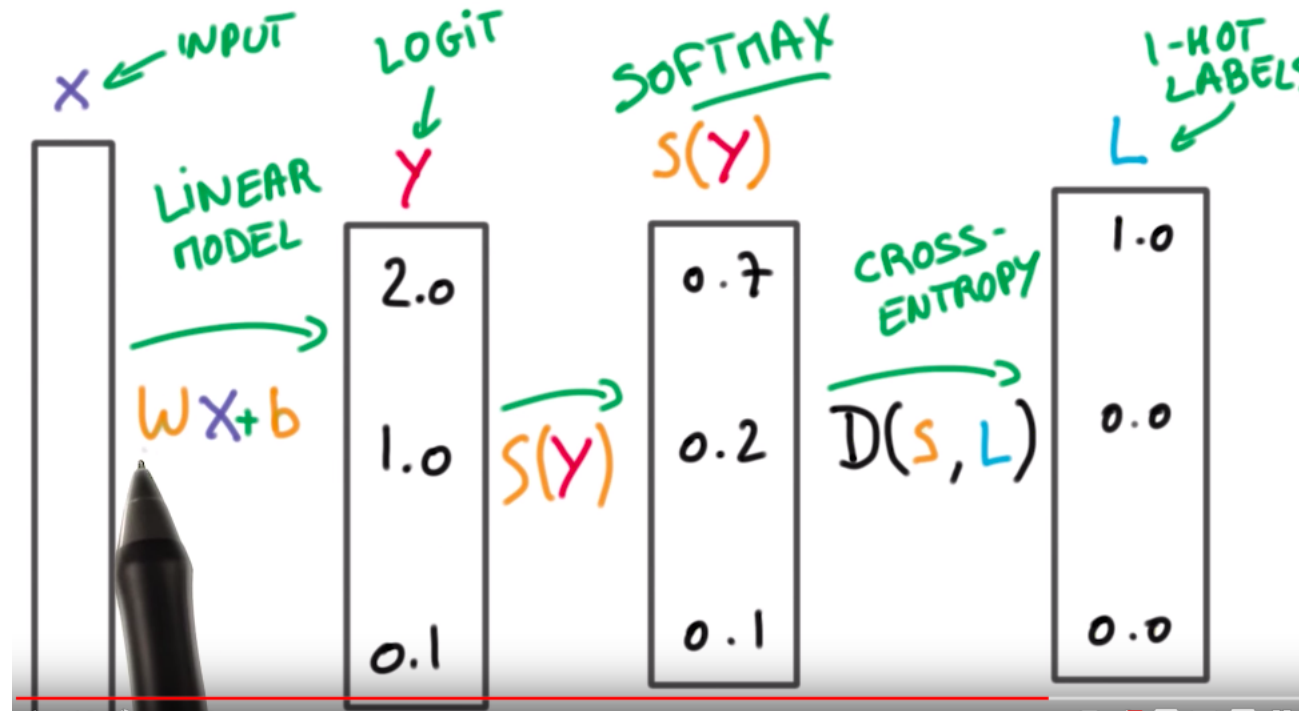
- The softmax function is often used in the final layer of a neural network-based classifier.
- All probabilities sum to one
- Often used with a [log loss](#) (or [cross-entropy](#)) cost function
- To solve a non-linear variant of multinomial logistic regression.

Activation Function: Softmax



<https://medium.com/data-science-bootcamp/understand-the-softmax-function-in-minutes-f3a59641e86d>

Loss Function: Cross Entropy

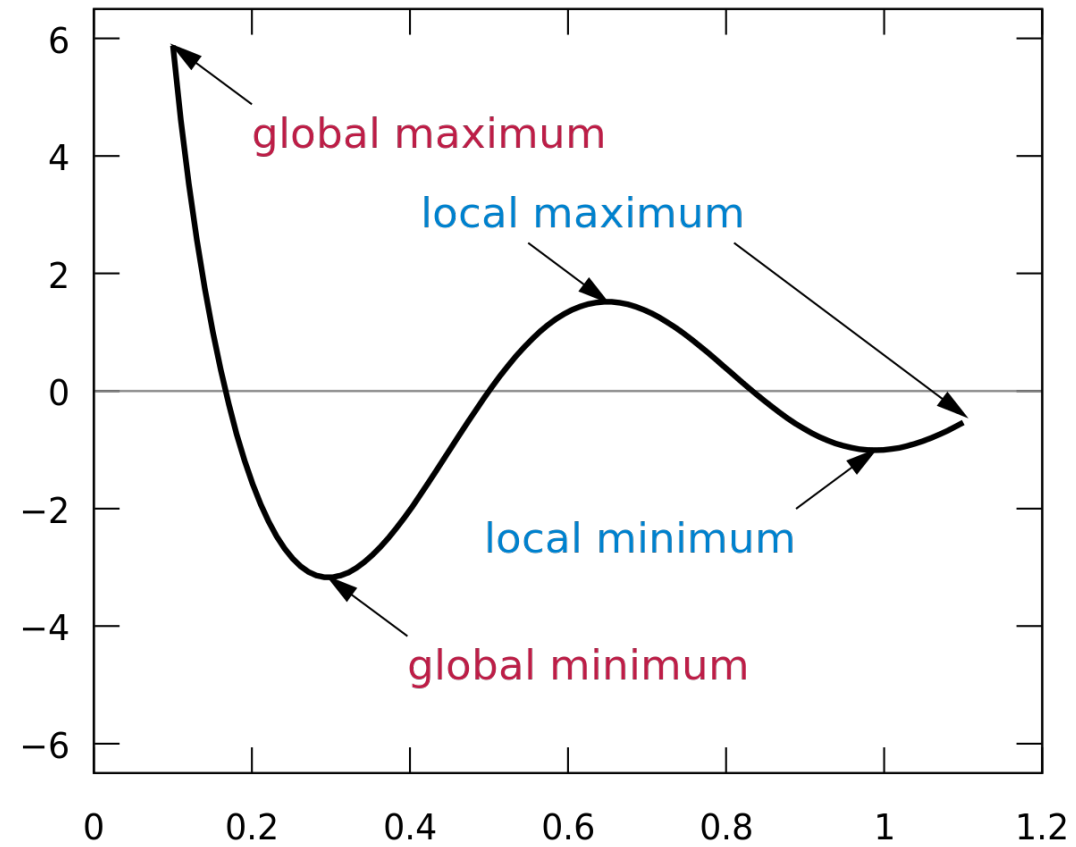


<https://medium.com/data-science-bootcamp/understand-cross-entropy-loss-in-minutes-9fb263caee9a>

Problems with training

- Initial weights: random means you can't predict
- Vanishing/exploding gradients
- Local minima

Local minima



Lab time

- To clone: from your terminal
 - >git clone <https://github.com/changsin/DeepLearning-101.git>
- Or use google colab to point to the git hub repository
- Git is an open source version control system
 - Github is a host service using git.