



AI Champ

Deep Learning Basics

Session - Oct 3 2020



Topics

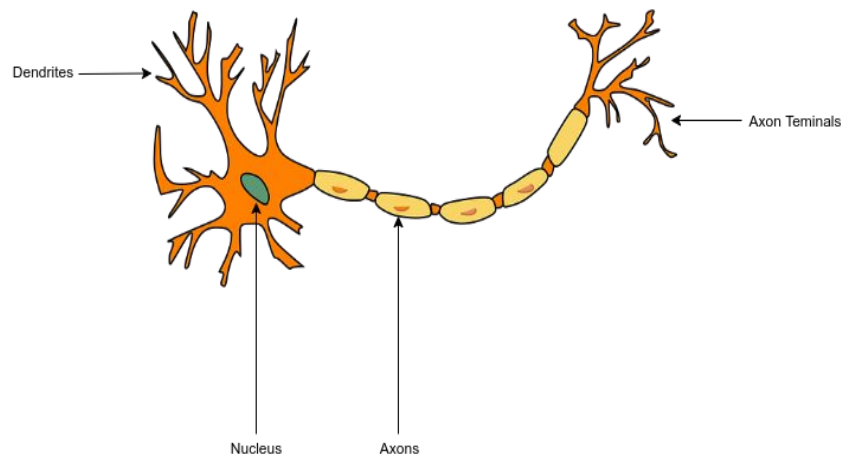
1. Deep Learning history and inspiration from the human brain.
2. An overview of mathematics will help you understand the language used in deep learning.
3. ANN and its components - Perceptron, Layers, Weight, Bias, Loss, Forward propagation, Back propagation, Linear and Nonlinear Functions, Activation Functions.
4. Understanding the general problems and some tips to train the models - Imbalanced Datasets, Overfitting, Underfitting, Regularization, Augmentation, Small Datasets, Transfer Learning, Vanishing Gradients, Exploding Gradients, Normalization.



Inspiration from Human Brain

Inspiration from Human Brain

1. Approximately 86 billion neurons in the human brain
2. Approximately 100 trillion connections between the neurons.
3. It takes some input signal, processes it and sends some output signal to other neurons.



Basic Maths for DL

Basic Maths for Deep Learning

1. Notations → <https://www.flickr.com/photos/95869671@N08/40544016221>
2. Scalars → Single number, can be arrays of numbers (magnitude only)
3. Vectors → Array of numbers arranged in order. It contains magnitude as well as direction.
4. Matrices → 2D array of numbers.
5. Tensors → Arrays with more than 2 dimensions. Basically, matrices with variable number of axes.

Basic Maths for Deep Learning

6. L^p Norm $\rightarrow \|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$
7. Matrix Transpose \rightarrow Interchange corresponding rows and columns.
8. Matrix Multiplication \rightarrow Only possible if dimensions of matrices are $P \times Q$ and $Q \times R$. Multiply Row items to corresponding Column items.
9. Equation of Hyperplane $\rightarrow \mathbf{Y} = \mathbf{W} * \mathbf{X} + \mathbf{B}$

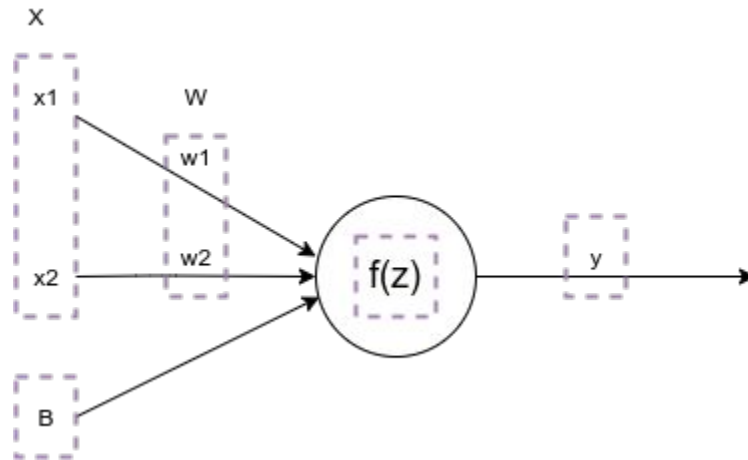


Artificial Neural Networks

ANN - Components

1. Perceptron
2. Layers
3. Weight
4. Bias
5. Loss
6. Forward Propagation
7. Backward Propagation
8. Linear and Nonlinear Functions
9. Activation Functions

Perceptron



$$z = W * X + B$$

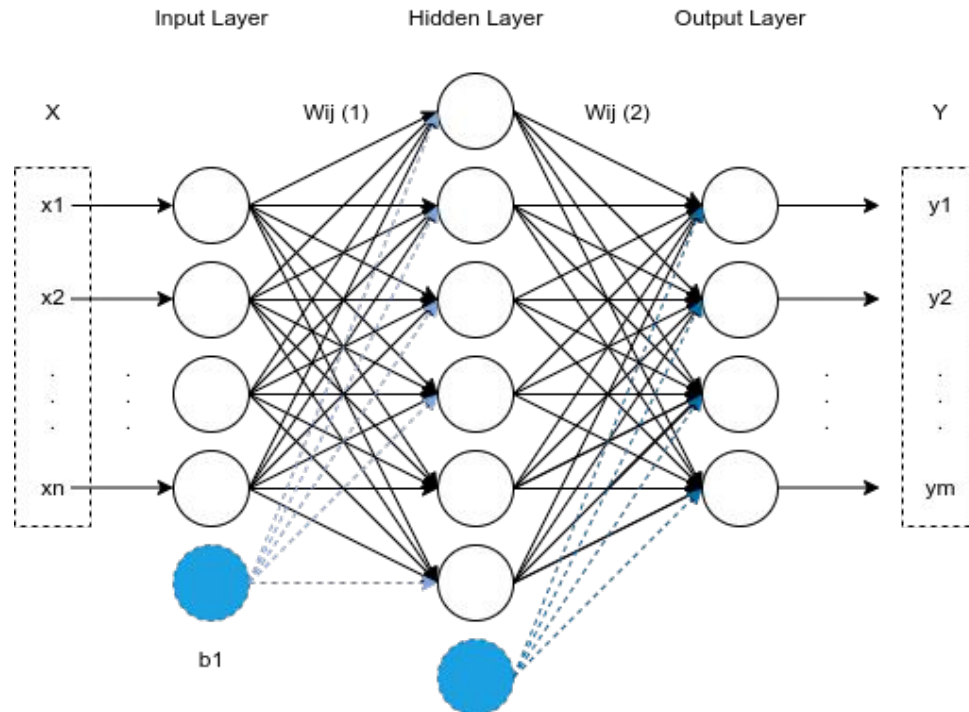
$$= x_1 * w_1 + x_2 * w_2 + B$$

$$y = f(z)$$

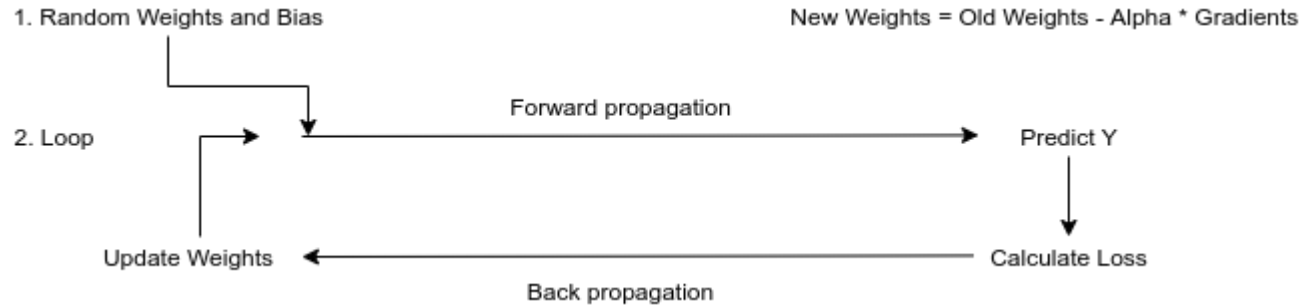
$$f(z) = \begin{cases} 0 & \text{if } z < T \\ 1 & \text{if } z \geq T \end{cases}$$

Where, T is threshold.

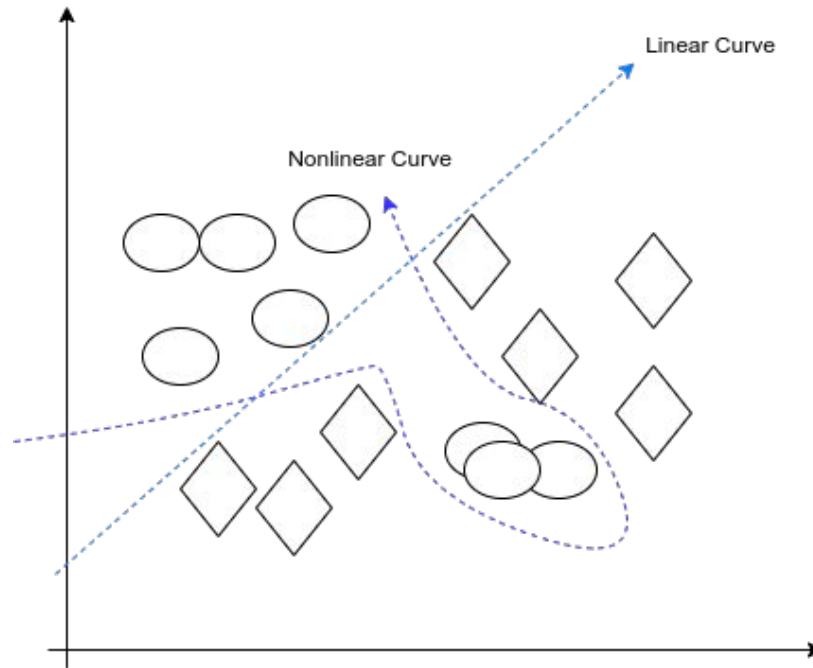
ANN



ANN - Training Process



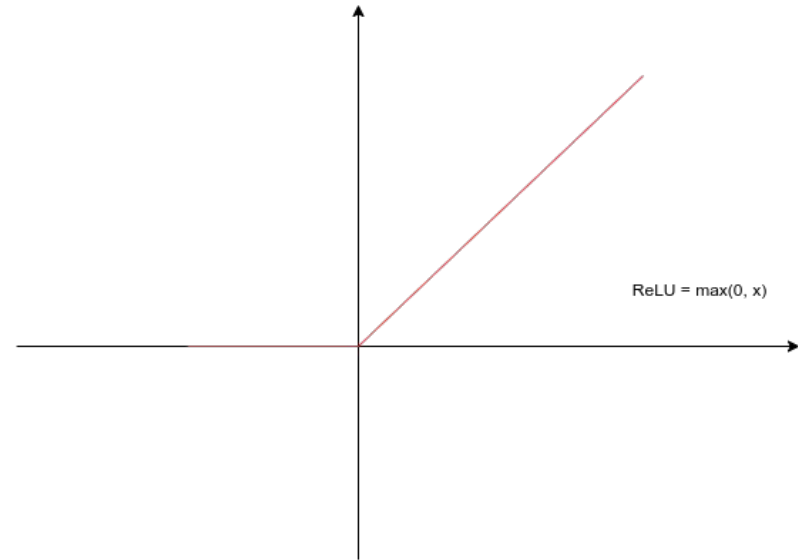
Linear and Nonlinear functions



Activation functions

Activation functions are nonlinear functions used to introduce nonlinearity in the deep learning models.

Some activation functions are - ReLU, Tanh, Sigmoid, Leaky ReLU



Resources to Look into

1. But what is a Neural Network?
https://www.youtube.com/watch?v=aircAruvnKk&ab_channel=3Blue1Brown
2. Gradient Descent
https://www.youtube.com/watch?v=lHZwWFHWa-w&ab_channel=3Blue1Brown



General Problems and Methods to Train the Model

General Problems and Methods to Train the Model

1. Imbalanced Datasets
2. Small Datasets
3. Augmentation
4. Overfitting
5. Underfitting
6. K-Fold Cross Validation
7. Regularization
8. Dropout
9. Vanishing Gradients
10. Exploding Gradients
11. Normalization
12. Weight Initialization
13. Transfer Learning

Dropout

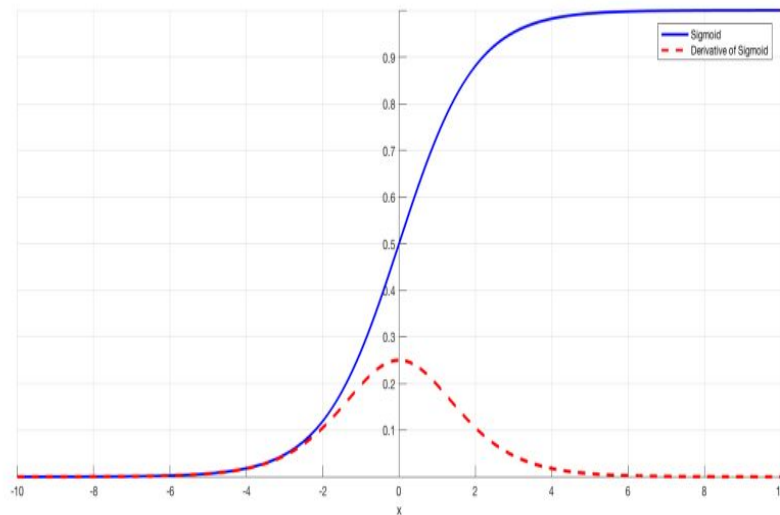
Randomly make given percentage of neurons output to 0.

For eg. Dropout(0.5) will randomly make 50% of neurons output to 0.

This is used to reduce overfitting.

Vanishing Gradient

Vanishing Gradient problem happens when we use certain activation functions (such as sigmoid) in model having large number of layers.



Exploding Gradient

Exploding Gradient problem happens when the gradients increase exponentially.

Model is not able to learn properly from the training data due to this.

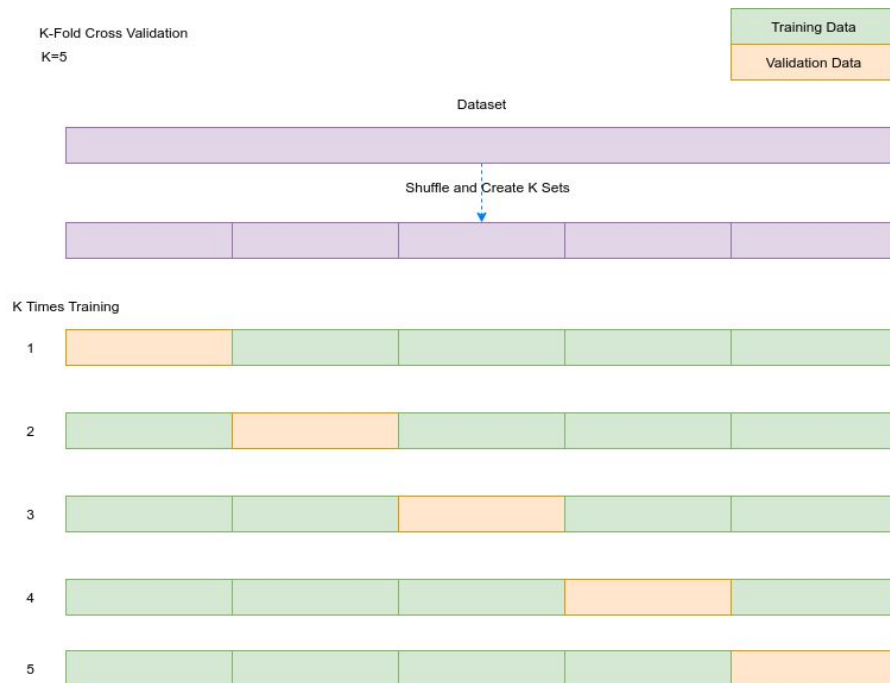
K-Fold Cross Validation

We do K-Fold Cross Validation to reduce overfitting.

The whole dataset is divided into K sets.

K-1 sets are used for training.
1 set is used for validation.

We repeat the training K times and make sure that every set is used as validation set.



Transfer Learning

What will happen if you train your model on a dataset and then use that trained model on different dataset and perform some fine-tuning so that it also learns the patterns in the new dataset?

- It will learn faster
- Accuracy also increases
- You require very less data on 2nd dataset



Week Work

Week Work

1. Create AND, OR, NOT gate using perceptron. Why can't we create for XOR?
2. CIFAR-10 Task with CNN
3. Do CIFAR with Transfer Learning - Use architecture of your choice.
4. Cat-Dog Classification - Use <https://www.kaggle.com/c/dogs-vs-cats/data> as your Test data. Don't use it for training. For the training data - you should gather the data from other sources - such as google images, bing images.
5. Draw Confusion Matrix for the above task and, calculate Precision, Recall, Accuracy, Sensitivity, Specificity, True Positive Rate, False Positive Rate.
6. Contribute to open source - Hacktoberfest - Get free t shirt!