**DEEP LEARNING**

* Deep learning is the subset of machine learning and machine learning is the subset of artificial intelligence.
* Inspired by the structure of the human brain, this structure is known as ANN(Artificial neural network).
* It refers to training the neural Networks.

**NEURAL NETWORKS:**

* Neural networks, particularly deep learning models, have gained significant attention in recent years due to their remarkable performance in various machine learning tasks.
* Neural networks are computational models inspired by the biological neural networks of the human brain.
* Components of neural networks:

A neural network consists of interconnected nodes called neurons, organized into layers. Typically, there are three types of layers: input layer, hidden layers, and output layer.



**Types of Neural Networks:**

**Feedforward Neural Networks (FNN):**

* Basic neural network where information flows in one direction, from input to output layer.
* Often used for simple classification and regression tasks.

**Convolutional Neural Networks (CNN):**

* Designed for processing grid-like data, such as images.
* Utilizes convolutional layers to extract spatial hierarchies of features.
* Widely used in image recognition, object detection, and image segmentation tasks.

**Recurrent Neural Networks (RNN):**

* Suitable for sequential data with temporal dependencies, such as time series or natural language.
* Contains loops that allow information to persist over time.
* Often used in language modeling, speech recognition, and time series prediction.

**Long Short-Term Memory Networks (LSTM):**

* A type of RNN designed to overcome the vanishing gradient problem.
* Includes memory cells and gates to control the flow of information, enabling better long-term dependencies modeling.
* Used in tasks requiring modeling of long-range dependencies, such as language translation and sentiment analysis.

**Neuron Operation:**

Each neuron computes a weighted sum of its inputs, applies an activation function to the result, and passes the output to the next layer.



**Training process of neural networks:**

**1.Forward Propagation**:

* During forward propagation, input data is fed into the neural network, and computations are performed layer by layer to generate predictions.
* Each neuron in a layer receives inputs from the previous layer, computes a weighted sum of these inputs, applies an activation function, and passes the result to the next layer.
* The process continues until the output layer is reached, and the final predictions are generated.
* The output of the neural network is compared with the ground truth labels to compute the loss function, which measures the discrepancy between the predicted and actual values.

**2.Backward Propagation:**

* Involves computing the gradients of the loss function with respect to the parameters of the neural network.
* Starting from the output layer, the gradients are propagated backward through the network using the chain rule of calculus.
* At each layer, the gradients are computed with respect to the activations and weights, allowing the network to understand how changes in these parameters affect the loss.
* The gradients are then used to update the parameters of the network (e.g., weights and biases) in the direction that minimizes the loss function, typically using optimization algorithms like gradient descent.

**Activation Functions:**

* Introduce non-linearity in the network.
* Decide whether to contribute to the next layer or not.
* Common activation functions include sigmoid, tanh, and rectified linear unit (ReLU), each serving different purposes in neural network architectures.

**Supervised Learning with Neural Networks:**

| Input(x) | Output(y) | Application |
| --- | --- | --- |
| Home features | Price | Real estate |
| Ad,user info | Click on Ad?(0/1) | Online Advertising |
| Image | Object(1,...1000) | Photo Tagging |
| Audio | Text transcript | Speech recognition |
| English | Chinese | Machine translation |
| Image,radar info | Position of other cars | Autonomous driving |

* - Supervised Learning: In supervised learning, the neural network learns from labeled training data, where each input is associated with a corresponding output.
* - Example: A common example of supervised learning is regression, where the neural network predicts a continuous value (e.g., house prices) based on input features (e.g., size, number of bedrooms).
* - Training Process: During training, the network adjusts its weights and biases iteratively to minimize the difference between predicted outputs and actual targets using techniques like gradient descent.
* Example:Housing Price Prediction
* - Dataset: Consider a dataset containing features such as house size, number of bedrooms, zip code, and neighborhood walkability.
* - Neural Network Architecture: Start with a simple neural network consisting of a single neuron for predicting house prices based on size.
* - Expansion: Expand the model to include additional features by stacking multiple neurons together, each processing a subset of input features.
* - Densely Connected Layers: The input layer and hidden layers are densely connected, allowing each neuron to receive input from all features.
* - Training and Prediction: Train the neural network using labeled data (X and Y pairs) to learn the relationship between input features and house prices. Once trained, the network can predict house prices for new input data.

**Deep Learning Growth**

• Data Availability: Large datasets like text, images, videos, and sensor data provide ample training examples for deep learning models.

• Computational Power: Advances in hardware, particularly GPUs and TPUs, have accelerated the training of deep learning models.

• Algorithmic Advances: Continuous improvements and refinements in deep learning algorithms have improved model architectures, regularization techniques, and optimization algorithms.

• Open-Source Ecosystem: Open-source deep learning frameworks like TensorFlow, PyTorch, and Keras provide developers with powerful building blocks for designing, training, and deploying deep learning models.

**TERMINOLOGIES IN NEURAL NETWORK**:

1. **Activation Function:**

* A nonlinear function applied to the weighted sum of inputs to introduce non-linearity into the network.

1. **Loss Function:**

* Quantify the deviation between the predicted output to the expected output.
* The error is computed for a single training network.
* The loss function evaluates how close a model's predictions are to the actual data. It directs the model to adjust its parameters to minimize this difference during training, improving its performance.

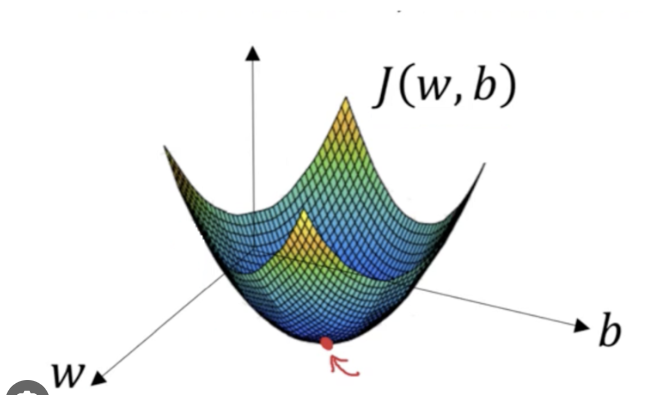


1. **Optimizers:**

* Optimizers are algorithms used during the training of machine learning models to minimize the cost or loss function
* Their primary role is to update the model's parameters iteratively to find the optimal values that minimize the cost function.

**Gradient Descent:**

* A fundamental optimization algorithm that updates the model's parameters in the direction of the steepest descent of the cost function.
* It iteratively updates the model's parameters in the direction of the steepest descent of the cost function with respect to those parameters.



* J is a convex function and J(w,b) measures how well the parameter performs on the entire training set.

**VISUAL REPRESENTATION OF GRADIENT DESCENT:** 

1. **Parameters and Hyperparameter:**

**Model Parameters:**

* Model parameters are the internal variables that a machine learning model learns from the training data in order to make predictions.
* It is estimated from data and not manually set.
* It includes weight(w) and bias(b).

**Model Hyperparameters:**

* Model hyperparameters are the configuration settings that define the architecture and behavior of a machine learning model.
* They are set prior to training and remain constant throughout the training process.

1. **Epoch,Batches,Batch size and Iteration:**

**Epoch:**

* An epoch refers to one complete pass through the entire training dataset during the training process.
* Training for multiple epochs allows the model to learn from the data multiple times and refine its parameters iteratively.

**Batch:**

* A batch is a subset of the training dataset that is used during each iteration of training.
* Instead of updating the model's parameters after processing the entire dataset ,training is done in smaller batches.

**Batch Size:**

* Batch size refers to the number of samples (data points) in each batch used during training.

**Iteration:**

* The number of iterations per epoch depends on the batch size and the size of the training dataset.
* It is equal to the total number of training samples divided by the batch size.

**1.Data Gathering:**

* Various sizes of dataset.

Amount of data needed=10\*no of parameters

* Quality of dataset.

**2.Data preprocessing**:

* Handle missing values, scale numerical features, encode categorical variables, engineer new features, perform data augmentation, address class imbalance, pad sequences, split the data, normalize input data, tokenize text data.

**3.Train the Model:**

* Train the model on the training data by feeding the input data through the model and adjusting the model's parameters (weights and biases) based on the computed loss.
* Monitor the model's performance on the validation set during training to detect overfitting and adjust hyperparameters as needed.

**4.Evaluate the Model:**

* Once training is complete, evaluate the trained model's performance on the test dataset to assess its generalization ability.
* Evaluation metrics such as accuracy, precision, recall, F1-score, or mean squared error, depending on the nature of the task.

**5.Optimizing:**

* Fine-tuning the model and hyperparameters based on the evaluation results to improve performance.
* Iterating on the model-building process by experimenting with different architectures, hyperparameters, and preprocessing techniques.