Basics of Multilayer Perceptron

A Multilayer Perceptron (MLP) is a type of artificial neural network (ANN) that consists of multiple layers of neurons (nodes). It is one of the simplest and most widely used architectures for supervised learning tasks, such as classification and regression.

Key Features of Multilayer Perceptron:

1. **Layered Structure**:
   1.1. Input Layer: Takes the input features of the data.
   1.2. Hidden Layers: One or more intermediate layers where computations are performed. Each neuron in these layers applies a weight to its inputs, adds a bias, and passes the result through an activation function.
   1.3. Output Layer: Provides the final prediction or classification.

2. **Fully Connected:**
   2.1. Each neuron in a layer is connected to every neuron in the subsequent layer. This structure is also called a feedforward network.
3. **Activation Function:**
   3.1. Non-linear activation functions like ReLU, sigmoid, or tanh are applied to the outputs of neurons to introduce non-linearity. This allows MLPs to model complex patterns.

4. **Training:**
   4.1. Training is typically done using backpropagation, a technique that calculates the gradient of the loss function with respect to each weight and bias using the chain rule.

4.2. An optimization algorithm like stochastic gradient descent (SGD) or Adam is used to update the parameters.

5. **Loss Function:**

5.1. The loss function quantifies the error between the predicted outputs and the actual labels. Examples include mean squared error (MSE) for regression and cross-entropy loss for classification.

Workflow of an MLP :

1. Input Data:
    a. Features are fed into the input layer.
2. Forward Propagation:
    a. Data passes through the network layer by layer.
    b. At each neuron, weights, biases, and the activation function are applied.
3. Output:
    a. Predictions are generated at the output layer.
4. Error Calculation:
    a. The difference between predictions and actual values is computed using the loss function.
5. Backpropagation:
    a. Gradients of the loss with respect to the parameters are computed and propagated backward to update weights and biases.
6. Repeat:
    a. Forward and backward propagation steps are repeated for multiple iterations (epochs) until the model converges.

Applications of Multilayer Perceptrons:

- Classification tasks (e.g., handwritten digit recognition, sentiment analysis)
- Regression tasks (e.g., predicting house prices)
- Function approximation
- Time series prediction
- Anomaly detection

Limitations:

- Not suitable for sequential data:
  - For temporal or sequential data, models like Recurrent Neural Networks (RNNs) or Transformers are preferred.
- Computationally intensive:
  - As the number of layers and neurons increases, training becomes computationally expensive.
- Overfitting:
  - MLPs with too many layers or neurons may overfit on small datasets.

There are 4 variants of MLP:

1. Univariate MLP Models
2. Multivariate MLP Models
3. Multi-Step MLP Models
4. Multivariate Multi-Step MLP Models

Summary of Classification Criteria:

Criteria	Types
Number of Layers	Single-Layer, Multilayer
Number of Outputs	Binary, Multi-Class, Multi-Label
Input-Output Relationship	Classification, Regression
Learning Paradigm	Supervised, Unsupervised, Semi-Supervised
Network Depth	Shallow, Deep
Activation Functions	Linear, Non-Linear
Optimization Methods	Standard Gradient Descent, Advanced Optimization