# Research Report on Neural Networks

## 1. Basic Architecture and Components of Neural Networks

Neural networks are the artificial computer concept that is based on the human brain. It makes decisions about input data after recognizing the patterns. The basic structure of the neural network contains the composition of neurons, layers, and activation functions.

> **Neurons:** The neuron is the fundamental unit of computation in a neural network. Each neuron takes in input and an activation function is applied to compute its output. In a neural network, neurons are associated in layers.

### Neural networks generically consist of layers:

• **Input layer:** This layer is the raw data's first point of contact with the neural network. Any neuron in this layer represents one feature in the input data.

• **Hidden Layers:** An intermediate layer, or more than one, between the input and the output layer. These include various neurons using non-linear transformations-a weighted sum followed by an activation function-on the input data. In fact, a neural network may have one to many hidden layers with each containing any number of neurons. This is part of what helps make the network complex because now the model can learn complicated patterns associated with complex tasks.

• **Output Layer:** The last layer of the network offers the predictions or classifications of the network. Of course, the structure depends on the nature of the task, being a regression or classification. It introduces non-linearity into the network, thus allowing it to model complex relationships. Some common ones are as follows:

> • **Sigmoid:** These maps input values between 0 and 1 and can be useful in binary classification.
>
> • **ReLU** (Rectified Linear Unit): This outputs the input value if it is positive; otherwise, it returns zero. ReLU helps prevent the vanishing gradient problem; hence, it finds wide application in deep networks.
>
> •**Tanh:** It returns the input values in a range from -1 to 1 and, in certain cases, provides better convergence properties compared to sigmoid.
>
> •**Softmax:** It is used in the output layer for multi-class classification problems, taking raw scores and turning them into probabilities.

## 2. Types of Neural Networks and Their Applications

Neural networks are available in various architectures and are suited to different types of tasks:

**Feedforward Neural Networks:**
The simplest form of neural networks, where it does not have any cycle formed by the interconnection of neurons. Feedforward neural networks do not contain feedback loops, and information moves in one direction from input to output. FNN is applied for image and speech recognition, and a simple classification problem.

**Convolutional Neural Networks-CNNs:**
They include special architectures to process gridded data, images being the most common. CNNs are convolutional layers used to detect features such as edges and textures, while the pooling layers reduce the dimensionality of the input features but retain the most important ones. CNNs find broad applications in image recognition, object detection, and video analysis.

**Recurrent Neural Networks:**
These are designed to process data in sequences, keeping a memory of what the previous inputs were by internal feedback connections. This architecture is useful in applications where data possesses temporal relations and in natural language processing. Variants such as Long Short-Term Memory and Gated Recurrent Units alleviate problems in propagating gradients for very long sequences. RNNs are used in applications such as language modelling, machine translation, and speech recognition.

**Generative Adversarial Networks:**
This class of models consists of two competing networks: a generator and a discriminator. GANs are used to synthesize new data samples from a given dataset-say, generating realistic images or translating low-resolution images into high-resolution ones.

**Transformer Networks:**
They assign weight to different parts of the input using self-attention mechanisms. Transformers have been the backbone of state-of-the-art architectures in natural language processing, from BERT and GPT, because they have effectively captured long-range dependencies and large-scale data.

## 3. Training of neural networks

The training process of neural network involves the optimization of its parameters so that the difference between its predicted and true outcomes is reduced.

**Backpropagation:**

This is the fundamental algorithm in the training of neural networks. Backpropagation is a way to calculate the gradient with respect to each weight of a loss function using the chain rule and propagate the error backwards through the network. The steps involve:

- **Forward Pass**: Input data are fed through the network to get the predicted values.

- **Loss Calculation**: Compare prediction with actual output using some loss function, like mean squared error for regression and cross-entropy loss for classification.

- **Backward Pass:** Compute gradients of the loss function w.r.t network parameter.

- **Weight Update:** Modify the parameters of the models based on some optimization techniques to minimize the loss.

### Optimization Techniques:

An update of weights assigned over the learning rate follows:

- **Stochastic Gradient Descent:**
  This is computationally efficient, using a small batch of data, though at the cost of fine-tuning learning rates.

- **Momentum:**
  This enhances SGD by taking into consideration the gradient in the past that helps the optimizer escape local minima and speeds up the process of convergence.

- **Adam (Adaptive Moment Estimation):**
  Adam provides the advantages of both momentum and adaptive learning rates; hence it is very popular for training most types of neural networks.

## Bibliography


1. Amazon Web Services (2024) *What is a Neural Network?* Available at: https://aws.amazon.com/what-is/neural-network/#:~:text=A%20neural%20network%20is%20a,that%20resembles%20the%20human%20brain. (Accessed: 16 August 2024).

2. Massachusetts Institute of Technology (MIT) (2017) *Explained: Neural Networks and Deep Learning*. Available at: https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414 (Accessed: 18 August 2024).

3. Neural Networks (2024) *Neural Networks*. Available at: https://www.neuralnetworks.com.au (Accessed: 21 August 2024).

4. Medium (2024) *Overview of a Neural Network's Learning Process*. Available at: https://medium.com/data-science-365/overview-of-a-neural-networks-learning-process-61690a502fa (Accessed: 21 August 2024).

5. MyGreatLearning (2024) *Types of Neural Networks*. Available at: https://www.mygreatlearning.com/blog/types-of-neural-networks/ (Accessed: 24 August 2024).

6. Dremio (2024) *Neural Network Architecture*. Available at: https://www.dremio.com/wiki/neural-network-architecture/#:~:text=The%20construct%20of%20a%20neural,to%20the%20task%20at%20hand. (Accessed: 25 August 2024).