“ An overview of neural networks ”

By Angus Wallbrink

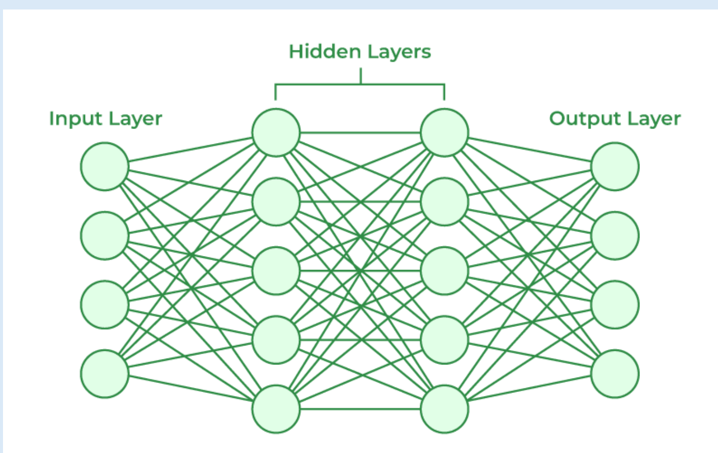
Understanding Neural Networks

Neural network, a type of deep learning, uses interconnected nodes and neurons in a layout alike to a human brain to process complex tasks like predictive modeling, image recognition, and solving problems in artificial intelligence. A neural network is an interconnected group of neurons, in Image 1.1 each node represents a neuron, and each line represents the output from one neuron travelling to the input of another.

Neural Network Architecture and Components

The output of a neuron is expressed by the formula [Output = Inputs* Weights + Bias]. The process occurs in the hidden layers, (layers between the input and output – See image 1.1) Weights and biases are parameters that determine how much influence a certain input has on the output. When a network is learning, these are adjusted to ensure there is minimal difference between the predicted output and the actual target.

Image 1.1:



The basic structure of a neural network consists of the input layer, hidden layers, and the output layer, all of which work together to process data in a way similar to the human brain. Here is a more in-depth breakdown of each layer:

- Input layer: The input layer receives data or inputs from the outside world e.g., an image, text, video etc. Then accepts it and passes it through to the rest of the network.
- Hidden layer: The hidden layer transforms the data by applying weights, biases, and activation functions, and through doing so extracts patterns and features of the data
- Output layer: Produces the final predictions or results.

There are also other components that play an important role in neural networks such as algorithms and functions. For example:

- Activation functions: A function that decides whether the neuron's input is important to the final output or not. This allows models to break away from linear relationships and process more complex non-linear relationships, allowing the neural network to model more complicated and sophisticated patterns. Common activation functions include:

Name	Formula	Purpose
Sigmoid	$S(x) = \frac{1}{1 + e^{-x}}$	The sigmoid function allows for values between 0 and 1. This is important when predicting the probability of events.
ReLU	f(x) = max (0, x)	ReLU introduces non-linearity, this allows neural networks to learn more complex relationships in data.
Tanh	$Tanh(x) = \frac{e^{x}-e^{x}}{e^{x}+e^{-x}}$	The tanh function produces values between –1 and 1, allowing the network to learn positive and negative patterns and relationships.

- Loss function: A loss function assesses how effective a network is by measuring how accurately a neural network models the training data. When training a model, you want to minimize this loss between predicted and target outputs.
- Optimization algorithms: These are closely related to loss functions as they aim to reduce losses. This is achieved by changing various parameters of a neural network such as weights and biases.

Different types and applications

Feedforward Neural Networks:

Feedforward Neural Networks (FNN) are the simplest form of neural networks and are characterized by their one directional flow of information – from input, through the hidden layers to the output. The straightforward design of FNN makes them well suited to applications like image recognition and simple pattern processing.

Convolutional Neural Networks:

Convolutional Neural Networks (CNN) are similar to FNNs in their omnidirectional flow of data however differ slightly in their architecture. CNNs have 3 main layers within their hidden layer:

- Convolutional layers: Apply filters to features present
- Pooling layers: Reduces the size of the feature map, cutting computation and memory use
- Fully connected layers: Combines the features learnt in the previous 2 layers to make the final predictions and classifications.

The use of these layers makes CNNs optimal for image-based predictions such as image recognition, object detection and video analysis.

Recurrent Neural Networks:

Recurrent Neural Networks (RNN) are a class of neural networks used for sequential data processing. Sequential data is data—such as words, sentences, or time-series data—where sequential components interrelate based on complex semantics and syntax rules. Unlike FNNs which only pass data through the network once, RNNs process data multiple times allowing the model to learn more complex patterns and relationships over time, this makes RNNs useful for applications such as natural language processing and speech recognition.

Training Processes and Techniques

Data preparation:

Data preparation is the process of making sure your data is clean, consistent and in a format that means the rest of the network can run effectively. There are 3 main components of data preparation:

- Normalization: Scaling data to a consistent range to ensure consistency
- Splitting: Dividing the dataset into training, validation, and test sets to evaluate the model's performance
- Augmentation: Adding training examples to improve adaptability

Backpropagation:

Backpropagation adjusts weights and biases after each data pass through the model to minimize loss, repeating the process iteratively. As a type of supervised learning, it requires a known output for each input to calculate the loss, this ensures accurate results however because of the reliance on datasets with known input-output pairs. This can be resource-intensive, as labeling data often involves time and effort.

Optimization Techniques:

Like backpropagation, optimization techniques minimize the difference between predicted and actual values. One prominent technique is gradient descent, which iteratively alters parameters by moving them in direction that most reduces error.

Bibliography

Jain, V. (2019, December 30). Everything you need to know about "Activation Functions" in Deep learning models. Medium; Towards Data Science.

https://towardsdatascience.com/everything-you-need-to-know-about-activation-functions-in-deep-learning-models-84ba9f82c253?gi=251425f0334d

AWS. (2023). What is a Neural Network? AI and ML Guide - AWS. Amazon Web Services, Inc. https://aws.amazon.com/whatis/neuralnetwork/#:~:text=A%20neural%20network%20is%20a

Baheti, P. (2021, July 8). The Essential Guide to Neural Network Architectures. Www.v7labs.com. https://www.v7labs.com/blog/neural-network-architectures-guide

The Role of Weights and Bias in Neural Networks. (2023, October 11). GeeksforGeeks. https://www.geeksforgeeks.org/the-role-of-weights-and-bias-in-neural-networks/

SHARMA, S. (2017, September 6). *Activation Functions in Neural Networks*. Medium; Towards Data Science. https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

Introduction to Loss Functions. (n.d.). DataRobot AI Platform. https://www.datarobot.com/blog/introduction-to-loss-functions/#:~:text=At%20its%20core%2C%20a%20loss

Pramoditha, R. (2022, February 1). *Overview of a Neural Network's Learning Process*. Data Science 365. https://medium.com/data-science-365/overview-of-a-neural-networks-learning-process-61690a502fa

Great Learning Team. (2020, April 29). *Types of Neural Networks and Definition of Neural Network*. GreatLearning. https://www.mygreatlearning.com/blog/types-of-neural-networks/