

Bishara Hodali

Neural Networks

In order to comprehend the function of deep learning in unsupervised tasks, it is necessary to first define a neural network in its broadest definition. Neural networks, which are built to imitate biological processes, are fundamentally the engines for modern artificial intelligence. A neural network is described by Amazon Web Services (AWS) as a "method in artificial intelligence that teaches computers to process data in a way that is inspired by the human brain."

These networks are structurally made up of layers of connected nodes, or "neurons." And Here is a brief explanation of the layers in an Artificial Neural Network (ANN) based on the GeeksforGeeks article:

1. Input Layer Role: This is the network's entry point. It transfers raw data—such as an image's pixels or a dataset's columns—to the layers that follow.

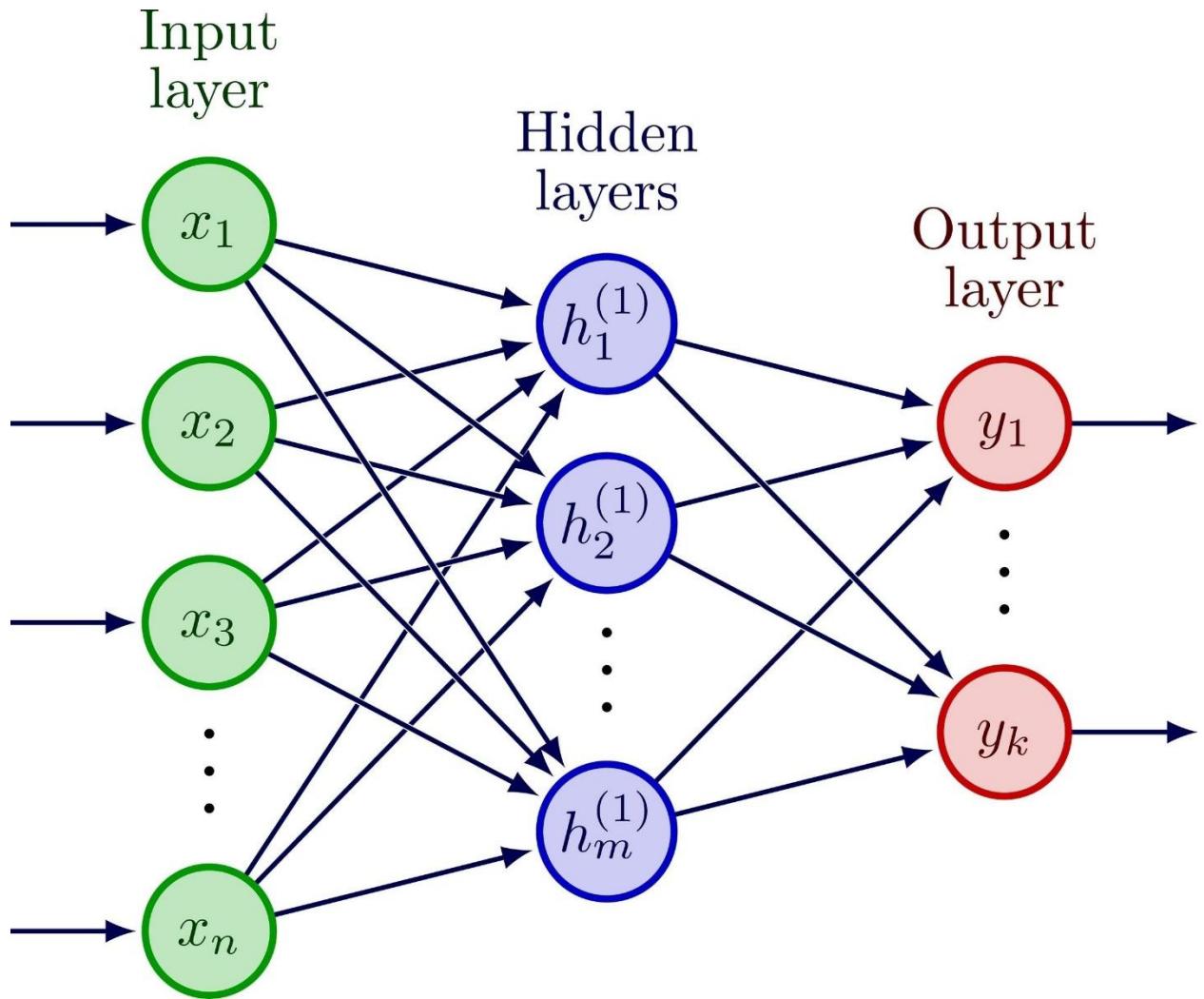
Key Feature: This layer merely buffers and sends the input signals to the buried levels; no calculation takes place here.

2. Hidden Layers: The real "learning" and processing takes place in these layers, which are situated between the input and output. A network is referred to as "Deep Learning" if it has one or more hidden layers.

Function: They use mathematical processes (weighted sums) and activation functions (like sigma or ReLU) to add non-linearity to extract features and patterns from the data.

The Hidden Layer Types can vary depending on the architecture. And these can be:

- Dense (Fully Connected) Layers : Every neuron connects to every neuron in the next layer.
- Convolutional (Convergent) Layers: Mostly used to identify geometric sequence in images.
- Recurrent (Repetitive) Layers: Applied to sequence data, such as text or periods of time.
- Pooling & Dropout Layers: used to avoid overfitting and minimize dimensions, respectively.



A diagram of a feedforward artificial neural network illustrating the flow of data through input, hidden, and output layers.

AWS states that "individual nodes can be simple, but when connected in a dense network, they solve complex problems." Supervised learning, in which the network is trained on labeled data (e.g., displaying computer images labeled "cat" until it learns to identify a cat), is traditionally the most well-known use of this design. This Data isn't always labeled, though. This leads us to the idea of clustering.

Clustering is a key method in unsupervised learning claims IBM. It involves "identifying distinct groups of data points" in which the machine must independently recognize patterns without the need for human assistance or pre-existing classifications. According to IBM, "clustering algorithms identify distinct groups of data points... such that data points

in the same group are more similar to other data points in the same group than those in other groups."

Neural Networks in Clustering on the Intersection

A shift in architecture is necessary to combine these two concepts—by utilizing a Neural Network, which is typically supervised, with Clustering, that is unsupervised. How does the network learn if there are no labels to fix it?

Standard feed-forward networks aren't ideally suited for this, as discussed in the Cross Validated (Stack Exchange) forum. Rather, certain designs are needed. The usage of automatic encoders is one popular strategy that was brought up in the conversation. One contributor points out that clustering can be accomplished by "training an autoencoder... and then clustering the data in the bottom layer." In this case, the clustering process is far more effective than attempting to group raw, high-dimensional data since the neural network learns to compress data (dimensionality reduction) into a dense representation.

However, there is a more direct "neural" approach to clustering that relies on competitive learning instead of error correction which is Self-Organizing Map (SOM).

Resources:

Amazon Web Services (AWS): <https://aws.amazon.com/what-is/neural-network/>

GeeksforGeeks: <https://www.geeksforgeeks.org/deep-learning/layers-in-artificial-neural-networks-ann/>

IBM: <https://www.ibm.com/think/topics/clustering>

Stats Stack Exchange (Cross Validated):

<https://stats.stackexchange.com/questions/140148/how-can-an-artificial-neural-network-ann-be-used-for-unsupervised-clustering>