Types of Deep Learning

[Deep](https://en.wikipedia.org/wiki/Deep_learning) learning offers human-like multi-layered processing in comparison with the shallow architecture. The basic idea of deep learning is to employ hierarchical processing using many layers of architecture. The architecture layers are arranged hierarchically. After several pre-training, each layer’s input goes to its adjacent layer. Most often, such pre-training of a selected layer executed in an unsupervised way. Deep learning follows a distributed approach to managing big data. The method assumes that the data gets generated considering numerous factors, different time, and various levels. Deep learning facilitates the arrangement and processing of the data into different layers according to its time (occurrence), its scale, or nature.

# Multi-layer perceptron

* This type of network is having more than 3 layers and it’s used to classify the data which is not linear
* These kinds of networks are fully connected with every node.
* These networks are extensively used for speech recognition and other machine learning technologies.

# Artificial Neural Network

* An artificial Neural Network is the component of a computing system designed in such a way that the human brain analyses and makes a decision. Ann is the building block of deep learning and solves the problem that seems impossible or very difficult by humans.
* Artificial neural networks work like a human brain. The human brain has billions of neurons, and each neuron is made up of a cell body that is responsible for computing information by carrying forward information towards hidden neurons and provide final Output.
* ANN initially in the training phase learns to identify patterns based on inputs given to the input layer. During this phase, the output of Ann compares with the actual output, and the difference between these two knows as an error.
* The aim is to minimize the error by adjusting the weight and bias of the interconnection which is known as backpropagation. With the process of backpropagation, the difference between the desired output and actual output produces the least error.
* ANN can be used to solve problems related to: Tabular data, Image data, Text data
* Advantages include:
  + Artificial Neural Network is capable of learning any nonlinear function. Hence, these networks are popularly known as Universal Function Approximators. ANNs have the capacity to learn weights that map any input to the output.
  + One of the main reasons behind universal approximation is the activation function. Activation functions introduce nonlinear properties to the network. This helps the network learn any complex relationship between input and output.
* Disadvantages include:
  + While solving an image classification problem using ANN, the first step is to convert a 2-dimensional image into a 1-dimensional vector prior to training the model. The number of trainable parameters increases drastically with an increase in the size of the image. Also, the ANN loses the spatial features of an image. Spatial features refer to the arrangement of the pixels in an image.
  + ANN cannot capture sequential information in the input data which is required for dealing with sequence data.

# Convolution Neural Network (CNN)

* CNN is one of the variations of the multilayer perceptron.
* CNN can contain more than 1 convolution layer and since it contains a convolution layer the network is very deep with fewer parameters.
* CNN is very effective for image recognition and identifying different image patterns.
* A convolutional neural network (CNN) is another variant of the feedforward multilayer perceptron. It is a type of feedforward neural network, where the individual neurons are ordered in a way that they respond to all overlapping regions in the visual area.
* Deep CNN works by consecutively modelling small pieces of information and combining them deeper in the network. One way to understand them is that the first layer will try to identify edges and form templates for edge detection. Then, the subsequent layers will try to combine them into simpler shapes and eventually into templates of different object positions, illumination, scales, etc. The final layers will match an input image with all the templates, and the final prediction is like a weighted sum of all of them. So, deep CNNs can model complex variations and behaviour, giving highly accurate predictions.
* ANN can be used to solve problems related to: Time Series data, Text data, Audio data
* Advantages include:
  + RNNs share the parameters across different time steps. This is popularly known as Parameter Sharing. This results in fewer parameters to train and decreases the computational cost.
  + RNN captures the sequential information present in the input data i.e., dependency between the words in the text while making predictions.

# Recurrent Neural Network

* RNN is a type of neural network where the output of a particular neuron is fed back as an input to the same node.
* This method helps the network to predict the output.
* This kind of network is useful in maintaining a small state of memory which is very useful for developing the chatbot
* This kind of network is used in chatbot development and text-to-speech technologies.
* The convolutional model works on a fixed number of inputs, generates a fix-sized vector as output with a predefined number of steps. The recurrent networks allow us to operate over sequences of vectors in input and output. In the case of recurrent neural network, the connection between units forms a directed cycle. Unlike the traditional neural network, the recurrent neural network input and output are not independent but related. Further, the recurrent neural network shares the standard parameters at every layer. One can train the recurrent network in a way that is like the traditional neural network using the backpropagation method.
* Here, calculation of gradient depends not on the current step but previous steps also. A variant called a bidirectional recurrent neural network is also used for many applications. The bidirectional neural network considers not only the previous but also the expected future output. In two-way and straightforward recurrent neural networks, deep learning can be achieved by introducing multiple hidden layers. Such deep networks provide higher learning capacity with lots of learning data. Speech, image processing, and natural language processing are some of the candidate areas where recurrent neural networks can be used.
* Advantages include:
  + CNN learns the filters automatically without mentioning it explicitly. These filters help in extracting the right and relevant features from the input data
  + [CNN](https://courses.analyticsvidhya.com/courses/convolutional-neural-networks-cnn-from-scratch?utm_source=blog&utm_medium=cnn-vs-rnn-vs-mlp-analyzing-3-types-of-neural-networks-in-deep-learning) captures the spatial features from an image. Spatial features refer to the arrangement of pixels and the relationship between them in an image. They help us in identifying the object accurately, the location of an object, as well as its relationship with other objects in an image.
  + CNN also follows the concept of parameter sharing. A single filter is applied across different parts of an input to produce a feature map.

Types of Machine Learning

# Supervised Machine Learning

* Supervised learning (SL) is the [machine learning](https://en.wikipedia.org/wiki/Machine_learning) task of learning a function that [maps](https://en.wikipedia.org/wiki/Map_(mathematics)) an input to an output based on example input-output pairs. It infers a function from labelled [training data](https://en.wikipedia.org/wiki/Training_set) consisting of a set of training examples. In supervised learning, each example is a pair consisting of an input object (typically a vector) and a desired output value (also called the supervisory signal). A supervised learning algorithm analyses the training data and produces an inferred function, which can be used for mapping new examples. An optimal scenario will allow for the algorithm to correctly determine the class labels for unseen instances. This requires the learning algorithm to generalize from the training data to unseen situations in a "reasonable" way ([inductive bias](https://en.wikipedia.org/wiki/Inductive_bias)). This statistical quality of an algorithm is measured through the so-called [generalization error](https://en.wikipedia.org/wiki/Generalization_error).
* Steps to follow:
  + Determine the type of training examples. Before doing anything else, the user should decide what kind of data is to be used as a training set. In the case of [handwriting analysis](https://en.wikipedia.org/wiki/Handwriting_analysis), for example, this might be a single handwritten character, an entire handwritten word, an entire sentence of handwriting or perhaps a full paragraph of handwriting.
  + Gather a training set. The training set needs to be representative of the real-world use of the function. Thus, a set of input objects is gathered, and corresponding outputs are also gathered, either from human experts or from measurements.
  + Determine the input feature representation of the learned function. The accuracy of the learned function depends strongly on how the input object is represented. Typically, the input object is transformed into a [feature vector](https://en.wikipedia.org/wiki/Feature_vector), which contains a number of features that are descriptive of the object. The number of features should not be too large; but should contain enough information to accurately predict the output.
  + Determine the structure of the learned function and corresponding learning algorithm. For example, the engineer may choose to use [support-vector machines](https://en.wikipedia.org/wiki/Support-vector_machine) or [decision trees](https://en.wikipedia.org/wiki/Decision_tree_learning).
  + Complete the design. Run the learning algorithm on the gathered training set. Some supervised learning algorithms require the user to determine certain control parameters. These parameters may be adjusted by optimizing performance on a subset (called a validation set) of the training set, or via [cross-validation](https://en.wikipedia.org/wiki/Cross-validation_(statistics)).
  + Evaluate the accuracy of the learned function. After parameter adjustment and learning, the performance of the resulting function should be measured on a test set that is separate from the training set.
* Examples Include:
  + Image Classification – The algorithm is drawn from feeding with labelled image data. An algorithm is trained, and it is expected that the algorithm classifies it correctly in the case of the new image.
  + Market Prediction – It is also called Regression. Historical business market data is fed to the computer. Then, with analysis and regression algorithm, the new price for the future is predicted depending on variables.

# Unsupervised Machine Learning

* Unsupervised learning is a type of [machine learning](https://en.wikipedia.org/wiki/Machine_learning) in which the algorithm is not provided with any pre-assigned labels or scores for the training data. As a result, unsupervised learning algorithms must first self-discover any naturally occurring patterns in that training data set. Common examples include [clustering](https://en.wikipedia.org/wiki/Cluster_analysis), where the algorithm automatically groups its training examples into categories with similar features, and [principal component analysis](https://en.wikipedia.org/wiki/Principal_component_analysis), where the algorithm finds ways to compress the training data set by identifying which features are most useful for discriminating between different training examples, and discarding the rest. This contrasts with [supervised learning](https://en.wikipedia.org/wiki/Supervised_learning) in which the training data include pre-assigned category labels (often by a human, or from the output of non-learning classification algorithm). Other intermediate levels in the supervision spectrum include [reinforcement learning](https://en.wikipedia.org/wiki/Reinforcement_learning), where only numerical scores are available for each training example instead of detailed tags, and [semi-supervised learning](https://en.wikipedia.org/wiki/Semi-supervised_learning) where only a portion of the training data have been tagged.
* Examples include:
  + Clustering – Data with similar traits are asked to group together by the algorithm; this grouping is called clusters. These prove helpful in the study of these groups, which can be applied to the entire data within a cluster more or less.
  + High Dimension Data – High dimension data is normally not easy to work with. With the help of unsupervised learning, visualization of high dimension data becomes possible.
  + Generative Models – Once your algorithm analyses and comes up with the probability distribution of the input, it can be used to generate new data. This proves to be very helpful in cases of missing data.
* Common families of algorithms used in unsupervised learning include: (1) clustering, (2) anomaly detection, (3) neural networks (note that not all neural networks are unsupervised; they can be trained by supervised, unsupervised, semi-supervised, or reinforcement methods), and (4) latent variable models.

# Reinforcement Machine Learning

* Reinforcement learning (RL) is an area of [machine learning](https://en.wikipedia.org/wiki/Machine_learning) concerned with how [intelligent agents](https://en.wikipedia.org/wiki/Intelligent_agent) ought to take [actions](https://en.wikipedia.org/wiki/Action_selection) in an environment in order to maximize the notion of cumulative reward. Reinforcement learning is one of three basic machine learning paradigms, alongside [supervised learning](https://en.wikipedia.org/wiki/Supervised_learning) and [unsupervised learning](https://en.wikipedia.org/wiki/Unsupervised_learning). Reinforcement learning differs from supervised learning in not needing labelled input/output pairs be presented, and in not needing sub-optimal actions to be explicitly corrected. Instead the focus is on finding a balance between exploration (of uncharted territory) and exploitation (of current knowledge).[[1]](https://en.wikipedia.org/wiki/Reinforcement_learning#cite_note-kaelbling-1)Partially supervised RL algorithms can combine the advantages of supervised and RL algorithms.
* This[type of machine learning algorithm](https://www.educba.com/types-of-machine-learning-algorithms/) uses the trial-and-error method to churn out output based on the highest efficiency of the function. The output is compared to find out errors and feedback fed back to the system to improve or maximize its performance. The model is provided with rewards which are basically feedback and punishments in its operations while performing a particular goal.
* Challenges of Reinforcement Machine Learning:
  + Reinforcement learning needs large datasets to make better benchmarks and decisions.
  + When the model’s complexity increases, reinforcement learning algorithms need more data to improve their decisions. That means the environments of the model may become more difficult to create reinforcement learning model.
  + The results of reinforcement learning models depend on the agent’s exploration of the environment and it brings limitation to the model. The agent takes actions according to the environment and its current states. If the environment changes constantly, making a good decision could be difficult.
  + Design of the reward structure of the model is another challenge for reinforcement learning. The agent uses the rewards and penalties to make a decision and perform its task. The way the agent is trained in the model is the key for the success.

# Semi-Supervised Machine Learning

* Semi-supervised learning is an approach to [machine learning](https://en.wikipedia.org/wiki/Machine_learning) that combines a small amount of [labelled data](https://en.wikipedia.org/wiki/Labeled_data) with a large amount of unlabelled data during training. Semi-supervised learning falls between [unsupervised learning](https://en.wikipedia.org/wiki/Unsupervised_learning) (with no labelled training data) and [supervised learning](https://en.wikipedia.org/wiki/Supervised_learning) (with only labelled training data). It is a special instance of [weak supervision](https://en.wikipedia.org/wiki/Weak_supervision).
* Systems using these models are seen to have improved learning accuracy