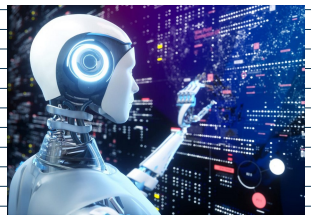
Terminologies

18 April 2023 20:57

- Artificial intelligence: it is simulation of the human intelligence process by machines. Especially in computer systems.
- Artificial intelligence is a field of science concerned with building computer and machines that can reason, learn, and act in such a way that would normally require human intelligence or that involves data whose scale exceeds what humans can analyze.
- artificial intelligence refers to computational tools that are able to substitute for human intelligence in the performance of certain tasks.



Types of artificial intelligence
1. reactive machines
2. limited memory
3. theory of mind
4. Self aware

Reactive machines: these are the oldest forms of AI systems that have extremely limited capability.
This machines cannot use previously gained experiences to inform their present actions.

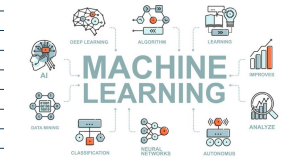
Limited theory: it is capable of learning from historical data to make decisions.

Theory of mind : machines will acquire true decision-making capabilities that are similar to humans.

Self aware : when machines can be aware of their own emotion, as well as the emotion of around them, they will have a level of consciousness and intelligence similar to human beings.

Example of AI : Alexa, face recognination, self-driving cars, text generation.

- **Machine learning:** it is an application of artificial intelligence that uses statical technique to enable computer to learn and make decisions without being explicitly programmed.
- Machine learning is a subset of artificial intelligence. It is focused on teaching computers to learn from data to improve with experience - instead of being explicitly programmed to do so.
- It uses computer that improve their efficiency automatically through experience.



Types of machine learning
supervised learning
Unsupervised learning
Semi-supervised learning
Reinforcement learning

Supervised learning: it is a class of problem that uses a model to learn the mapping between the input and target variables.

Unsupervised learning : it is an problem the model tries to learn by itself and recognize patterns and extract the relationships among the data.

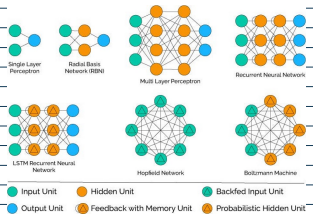
Semi-supervised learning: it is the combination of supervised and unsupervised learning.

Reinforcement learning: it is the type of problem where there is an agent and is operating in an environment based on the feedback or reward given

Neural networks: a method in artificial intelligence that teaches computers to process data in way tat is inspired by human brain.

A type machine leaning model that mimic the human brain or have a brain-like structure

It is a series of algorithms that endeavors to recognize underlying relationship in a set of data through a process that mimics the way human brain operates.



Types of neural network:

Artificial neural network: composed of a collection of connected nodes that takes input or set of inputs and return output.

Convolutional neural network: it is use the mathematical operations like convolution insted of general matrix multiplication in at lead one thir layer number.

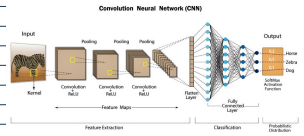
Recurrent neural network: it is connections between nodes dorm a diagraph along a temporal sequance allowing them to use their internal memory to process variable length sequences of input because of this chasterstic of ANN data edxeptional at landing sequence data like execution or audio recogination.

Examples: facial recognition, decision making

Convolutional neural network: it is the subset of artificial neural networks which we are used to different data types. The CNN is mainly used for image recognition, pattern recognition.

It is the type of neural network architecture commonly used in computer vision.

CNNs are particularly useful for finding patterns image to recognize objects, classes and categories.



CNN layers:

Convolutional layer : the majority of computation happen in convolutional layer which is the core building block for CNN. The second convolutional layer follow the initial layer. The process of convolution involves filter inside the layer moving across respective the fields of image. Checking features is present in the image.

Pooling layer: the pooling layer is also sweeps the kernel or filter across the image input pooling layer reduce the number of parameters in input image and also output. This layer complexity and improves the efficiency.

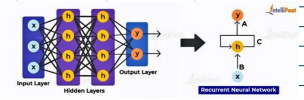
Fully connected layer : it is where image classification happens in the CNN based on features are extracted in the previous layer fully connected means all the inputs or nodes from one layer are connected to every activation unit or node of the next layer.

Example: image recoginition.

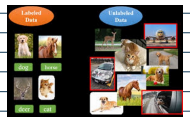
Recurrent neural networks: it is powerful and robust type of neural networks, and belongs to most promising algorithms in use because it is only one with in internal memory(it will remembers the important thing of input).

It is from the multilayer perceptron is that a recurrent neural network is built to handle inputs that represent a sequence, it handle output sequence.

It is act like a chain. The compution performed at each rime step, depends on the previous computation.



Labeled data: any data which has a charateristic, category, or attributes assigned to it can be referred to as labeled data.
Eg: photo of cat, the height of a human.



Data labeling: it is defined as process of identifying raw-like text, pdf, files, images and classifying models to learn from it.
Labeling helps the machine learning model identify the attributes of the data to analyse and make predictions.

Ways of labelling data:

Internal sourcing: large companies who have a dedicated in-house data science team can engage internal resources to label raw data.

Script labelling: data scientists can write code/scripts and run them to automatically annotate data. This reduces human intervention to a certain extent, but not completely.

Agencies: in case the organization dose not have sufficient bandwidth, it can outsource data labelling to agencies that specialize in such tasks.

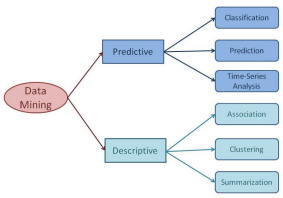
Unlabeled data: any data that does not have any labels specifying ints characteristics, identity, classification, or properties can be considered unlabeled data.
Eg: photos, videos

Data mining: data mining is the process of searching and analyzing a large batch of raw data in order to identify patterns and extract useful information.

Its functionalities are to perceive the various forms of patterns to be identified in data mining activities.

Data mining is the process of storing through large data sets to identify and relationships.

Data mining tasks:



Predictive data ming tasks come up with a model from the available data set that is helpful in predicting future values of another data set of interest.

Classification: it is used to create data structure classes, as the models is used to classify new instances whose classification is not understood.

Prediction: to detect the inaccessible data, it uses regression analysis and detects the missing numeric values in the data.

Time-seires analysis: it is the sequence of events where the next event is determined by one or more of the preceding events. Time series reflects the process being measured and there are certain components the affect the behavior of a process.

Descriptive data mining tasks usually finds data describing patterns and comes up with new, significant information from the available dataset.

Association: it discovers the association or connection among a set of items association identifies the relationships between objects.

Clustering: it is used to identify data objects that are similar to one another.

Summerization: it is the generalization of data. A set of relevant data is summarized which result

supervised and unsupervised learning.

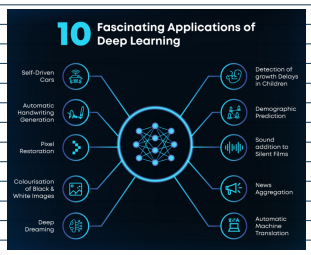
Reinforcement learning: it is the type of problem where there is an agent and is operating in an environment based on the feedback or reward given to the agent by the environment in which it is operating.

Example for machine learning: chatboat, google transaction, prediction.

Deep learning: it is a subset of machine learning, which is essentially a neural network with three or more layers.

It is artificial neural networks to perform sophisticated computations on large amounts of data. It is the type of machine learning that works based on the structure and function of the human brain.

It is method in artificial intelligence that teaches computers to process data in way that is inspired by the human brain.



types of Deep learning:

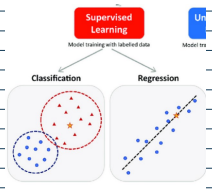
feed forward neural network: it is also known as artificial neural network which ensures that nodes do not form a cycle. It is fully connected which means that every neuron in layer is connected with all other neurons.
Recurrent neural network: it is perfect for time-related data and they are used in time series forcasting.
Convolutional neural network: it is special type of neural network mainly used for image classification oltering of image and object recognition.
Autoencoders : these are mostly used as an unsupervised algorithm and its main use-case is dimenionality reduction and compression.
It consists of encoder and decoder.

Examples of deep learning: virtual assistants, facial recognition, image and video recognition

Supervised learning: it is a subset of artificial intelligence and machine learning. It is also known as supervised machine learning. And it is defined by its ability to train algorithms to categorize data and predict outcomes accurately.

Supervised learning is the training machine learning model using training data is known as supervised learning.

Supervised learning involves machine learning algorithms that learn under the presence of a supervisor.



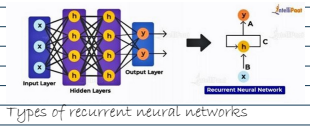
Types of supervised learning:

Classification
Regression

Classification: in classification, the out variable is a categorial variable, and the goal is to predict the class or category to which a new data point belongs.

Regression: in regression, the output variable is a continous variable, and the goal is to predict the value of the output variable based on the input variable.

Example: text categorization, face detection, spam detection, stock price predictions



Types of recurrent neural networks
One-one
One-many
Many-one
Many-many

Example: siri and google's voice search.

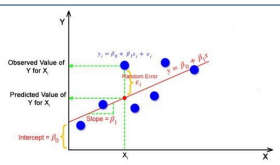
Deep neural network: it is an ANN with multiple hideen layers between the input and output layers. DNNs can complex non-linear relationships.

It is a sub-fieald of machine learning in artificial intlligence. That deals with algorithms inspired from the biological structure and functioning of a brain to aid machines with intelligence.

Linear regression: it tries to apply relations that independent variable data points the relation is usually a straight line that best fits.

It is defined as an algorithm that provides a linear relationship between an independent variable to predict the outcome of future events.

It is a supervised learning algorithm that simulates a mathematical relationship between variables and makes predictions for continuous or numeric variables



Formula: $Y = m * X + b$

X - dependent variable
Y - independent variable

$$Y(x) = P0 + P1 * X$$

Y - output variable represents the continous value that the model their to predict

X - input variable X is the feature independent variable in statistics.

P0 - y-axis intercept

P1 - the regression coeffient or scale factor is the equivalent of the slope of the best-fit straight line

Pi - wights (in general)

Types of linear regression

Simple linear regression: it revels the correlation between a dependent and independent variables.

Multi-linear regression: it establishes the relationship between independent variables and the corrseponding dependent variable.
Here independent variables can be either contineous or categorical.

Logistic regression: involves one dependent dichotomius variable(0, 1 values) and one independent variable.

Multinomial logistic regression: it performed when the dependent variable is nominal with more than two values.
It specifies the relationship between one dependent nominal variable and more contineous level (interval, ratio, or dichotomus) independent variable.

Logistic regressions: it is used when the dependent variable(target) is categorical.

Clustering: it is used to identify data objects that are similar to one another.

Summerization: it is the generalization of data. A set of relevant data is summarized which result in a smaller set that gives aggregated information of the data.

Structured and unstructured data:

Structured data: it consists of clearly defined data types with patterns that make easily searchable,

It is highly specific and is stored in predefined format.

It data that has been predefined and formatted toa set structure before being placed in data storage.

Unstructred data: it is data stored in its native format and not processes until used, which is known as schema-on-read.

Unstructred data is a complication of many varied types advantage of schema-on-write

Data prepossessing: it is a step in the data mining and data analysis process that takes raw data and transforms it into a format that can be understood and analysed by computer and machine learning.

It includes the steps we need to follow to transform or encode data so that it may be easily parases by the machine.

Data preprocessing is an important task. Is is data mining technique that transforms raw data into a more understandable, useful and efficient format.

Task in data processing:

Data cleaning: it is also known as scrubbing. This task involves filling of missing values, smoothing or removing noisy data and outliers along with resolving inconsistencies.

Data integration: this task involves integrating data from multiple sources such as database, data cubes, files, etc.

Data transformation: this involves normalisation and aggregation of data according to the needs of the data set

Data reduction: during this step data is reduces. The number of records or the number of attributes or dimensions can be reduced.

Data discretization: it is considered as a part of reduction. The numerical attributes are replaced with nominal ones.

Language model: it is the task of predicting the next word or character in a document.
This technique can be used to train language models that can further be applied to a wide range of natural language tasks like text generation, text classification, and question answering.

It uses machine learning to conduct a probability distribution over words used to predict the most likely next word in sentence based on the previous entry

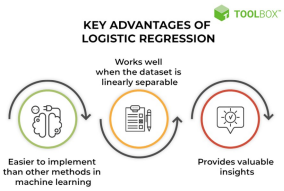
N-gram language models: a simple probilistic lanuage model is constructed by calculating n- gram probabilities. An n-gram is an a word sequence, n begin an integer grater than zero. N-grams's last word follows a praticular n-1 gram (leaving out the last word).

Neural language models: it is the based on the language models ease the sparsity problem by the way they encode inputs. Word embedding layers create an arbitary sized vector of each word that incorporates semantic relationships as well.

continous variable, and the goal is to predict the value of the output variable based on the input variable.

Example: text categorization, face detection, spam detection, stock price predictions.

variable (target) is categorical.



Types of logistic regression:

Binary logistic regression: the categorical response has only two 2 possible outcomes.

Multinomial logistic regression: three or more categories without ordering.

Ordinal logistic regression: three or more categories with ordering.

Unsupervised learning: unsupervised learning is a machine learning technique in which the users do not need to supervised the mosel. Instead, it aloows the model to work on its own to discover patterns and information that was previously undetected. It deals with unlabelled data.

It is the training the machine learning model using of unlabled data. Is known as unsupervised learning.

The main idea to learn under supervision where the supervision signal is named as target value or label.

Types of unsupervised learning:

Clustering: it is deals with finiding a structure or pattern in a collection of uncategorised data.

Association: association rules allow you to establish associations data objects inside large database.

Example : finding customer segments, pattern recognition,

Feedforward neural networks: it is the simplest type of artificial neural network which has lots of applications in machine learning.

It is also known as feedforward networks or multilayer preceptions. They form the basis of many important neural networks being used in the recent times, CNN, RNNs.

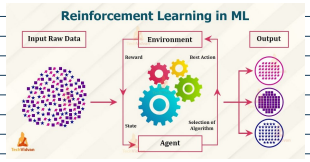
It is a single-layer perceptron in its most fundamental form.

It is the basic neural networks. Information is passed through several input nodes in one direction until it reaches the output node.

Reinforcement learning: it can be understood as a feedback-based machine learning algorithm or technique. In which an agent kept trying to learn within an environment through looking at it outputs or results.

It is the closet to human learning as digital systems and machines can get. Through this training, machine learning model can be taught to follow instruction, conduct test, operate and ect.

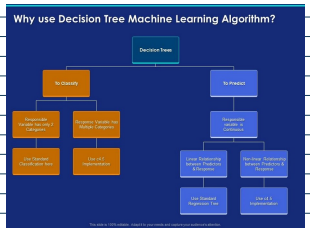
It is the type of machine larning that enables a computer system to learn how to make choices by being rewarded for its sucesses.



Decision tree: it is the type of supervised machine learning where data continuosly split according to the certain parameter.

It is create a training model that can use to predict class or value of the target variable by learning simple decision rule from training data.

It is used to categorise or make predictions based on previous set of questions were answered .



Types of reinforcement learning:

Positive reinforcement: positive reinforcement learning gives a positive impact on the action, which is taken by the agent, and increases the two factores

- Strength
- Frequency

Negative reinforcement: it can be the event, which aims to strengthen the behaviour that occurs due to negative condition which have to be stopped or avioded. It is defined to have minimum performance.

Example: marketing and adversising, gaming, automated robots.

Types of decision tree

Categorical variable decision tree : decision tree has a categorical target variable.

Continuous variable decision tree: decision tree has continuous Target variable

Regression: it is the method to determine the statistical relationship between a dependent variable and one more independent variables.

Regression analysis helps to predict the effect of the independent variable on the dependent one.

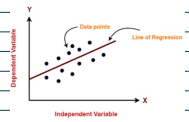
Random forest: it is a machine learning technique that's used to solve regression and classification problems.

It algorithm consists of many decision tress. The forest generated by the random forest algorithm is trained through bagging or bootstrap aggregating.

Random forest is a flexible, easy-to-use machine learning algorithm that produces even without hyper-parameter tuning, a great result most of the time.

variable and one more independent variables.

Regression analysis helps to predict the effect of the independent variable on the dependent one.



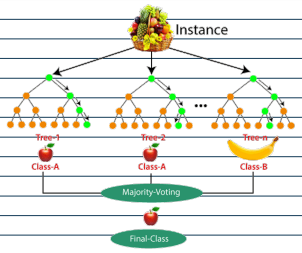
Types of regression

Linear regression: the dependent variable is continuous, or discrete and nature of relationship is linear. This relationship is known as line equation the best approximates all the individual data points.

Non-linear regression: if the regression is not linear and is in some other formm, then the regression is said to be non-linear regression.

Example: it can be used to predict the relationship between reckless driving and total number of road accidents caused by a driver.

Random forest is a flexible, easy-to-use machine learning algorithm that produces even without hyper-parameter tuning, a great result most of the time.



Congnitive science: it attempts to decode human thought process and underlying information processing mechanisms.

Cognitive science refers to the field of study that interface multiple dicipline such as neuroscience, computer science, psychology, artificial intelligence, philosophy, liguistics, and anthropology to understand congitive functioning of the human mind and the underlying mental process. It focuses on comprehending the nature of the human mind and how it uses mental representations to relize, process, transform, and manipulate knowledge.