AI.4.educators

AI.4.educators
Educating Educators on Artificial Intelligence (AI) – development of an AI training material and an AI educational program for educators

Project No: 2021-1-EL01-KA210-ADU-000034976

AI Practical Roadmap

Co-funded by the European Union

Table of Contents

Introduction to AI Practical Roadmap

The AI Practical Roadmap presents:

- The history of AI and how all started
- The most common ML Algorithms
- How Neural Networks works
- AI Applications in different sectors
- Examples of AI models

This information helps to understand what AI is and how can be used today.

History of AI- Birth of AI

- Warren McCulloch & Walter Pitts (1943): ANN with on-off neurons
    - Neurons triggered by sufficient #neighbors
    - Showed that any computable function computable with some network like this
    - Logical connectives implementable this way
    - Donald Hebb’s 1949 learning rule

- Two graduate students in the Princeton mathematics department, Marvin Minsky and Dean Edmonds, built the first neural network computer in 1951 called SNARC.

History of AI- Birth of AI

- Alan Turing: He invented the Turing Test, designed to determine if a computer system can be called an artificial intelligence or not, based on whether it can fool a human into thinking it is human too.

- U.S. researchers interested in automata theory, neural nets, and the study of intelligence were brought together in a workshop at Dartmouth in the summer of 1956 where John McCarthy suggested the name for the field as “artificial intelligence.”

History of AI- Early enthusiasm, great expectations

- Starting in 1952. Arthur Samuel wrote a series of programs for checkers (draughts) that eventually learned to play at a strong amateur level.

- McCarthy in 1958 clearly marked a high level language LISP, a dominant AI programming language. Furthermore, McCarthy and others at MIT invented time sharing at the same period. Also in 1958, McCarthy released a paper entitled Programs with Common Sense, in which he described the Advice Taker, a hypothetical program that can be seen as the first complete AI system.

History of AI- Early enthusiasm, great expectations

- Hebb's learning methods were enhanced by Bemie Widrow (Widrow and Hoff, 1960; Widrow. 1962), who called his networks adelines, and by Frank Rosenblatt (1962) with his perceptrons. Rosenblatt proved the perceptron convergence theorem.

History of AI- A Dose of Reality

- In 1966 the failure of machine translation project brought an end to the US government's funding of the project. • Minsky and Papert's book: 'Perceptrons' (1969) proved that. although perceptrons (a simple form of neural network) could be shown to learn anything they were capable of representing, they could represent very little.

- In 1973 Lighthill report entailed cutting of British funding to AI research in most of the universities in the Great Britain.

History of AI- Knowledge-based Systems

- The DENDRAL program (1969) solved the problem of inferring molecular structure from the information provided by a mass spectrometer.

- MYCIN was developed in mid 1970s at Stanford that diagnosed blood infections.

History of AI- AI becomes an industry

- The first successful commercial expert system R1 began operation at the Digital Equipment Corporation (McDermott, 1982)
- Nearly every major U.S. corporation had its own AI group and was either using or investigating expert systems.
- In 1981, the Japanese announced the "Fifth Generation" project, a 10-year plan to build intelligent computers running Prolog. • United States formed the Microelectronics and Computer Technology Corporation (MCC) as a research consortium.

History of AI- AI becomes an industry

- Alvey report reinstated the funding that was cut by the Lighthill report.
- The AI industry boomed from a few million dollars in 1980 to billions of dollars in 1988. Soon after that came a period called the "AI Winter" in which many companies suffered as they failed to deliver on extravagant promises.

History of AI- The return of neural networks

- In the mid-1980s at least four different groups reinvented the back-propagation learning algorithm first found in 1969 by Bryson and Ho. AI becomes a Science (1987-present)

- Judea Pearl's (1988) Probabilistic Reasoning in Intelligent Systems led to a new acceptance of probability and decision theory in AI.

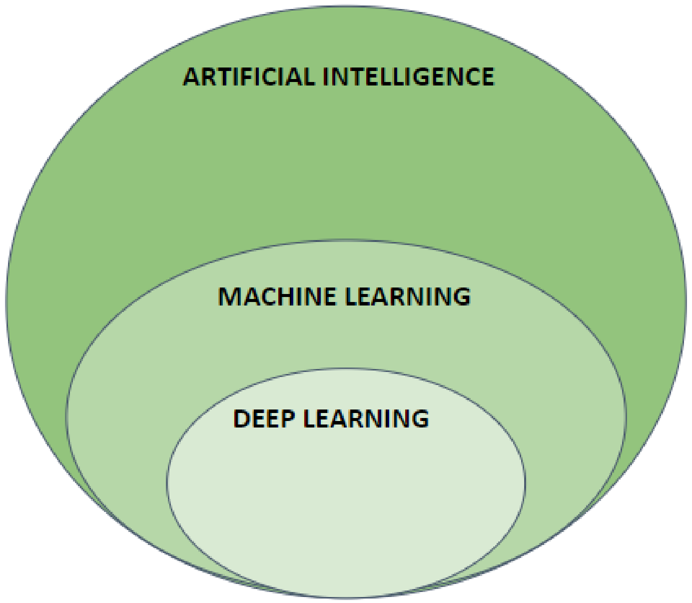
History of AI - The emergence of intelligent agents

- The work of Allen Newell, John Laird, and Paul Rosenbloom on SOAR (Newell. 1990: Laird el al., 1987) is the best-known example of a complete agent architecture.

- AI technologies underlie many Internet tools, such as search engines, recommender systems, and Web site construction systems

Machine Learning

Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

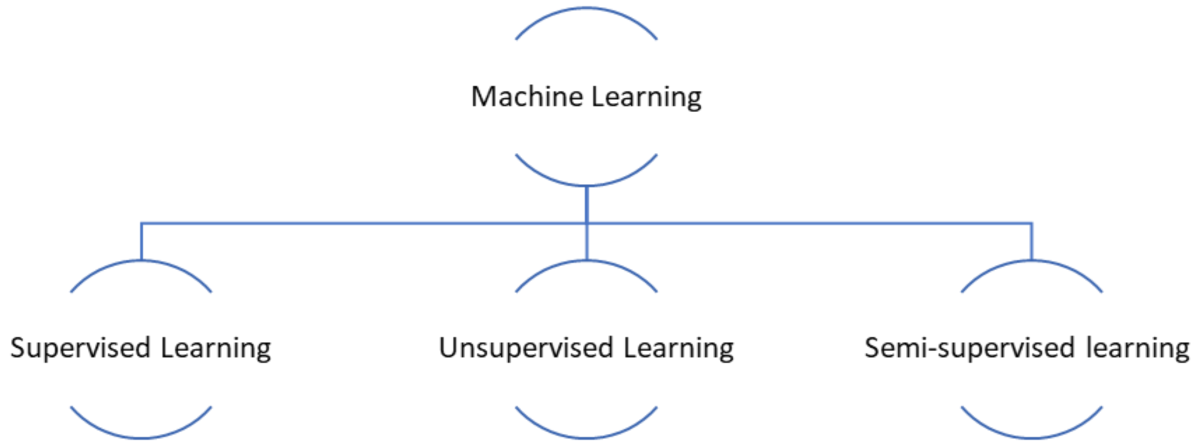


Machine Learning

Machine learning is an prominent component of the developing field of data science. By the use of statistical methods, algorithms are trained to make classifications or predictions, and to reveal key insights in data mining projects. These insights subsequently drive decision making within applications and businesses, perfectly impacting key growth metrics.

Machine Learning

With the continuously expansion of big data, the market demand for data scientists will increase. They will be required to help identify the most relevant business questions and the data to reply them. Machine learning models fall into three primary categories.



Supervised Learning

Supervised learning uses a training set to teach models to yield the desired output. This training dataset contains inputs and accurate outputs, which allow the model to learn over time. The algorithm measures its accuracy through the loss function, adjusting until the error has been sufficiently minimized.

Supervised learning can be separated into two types of problems when data mining—classification and regression:

Supervised Learning

- ***Classification*** uses an algorithm to accurately assign test data into specific categories. It recognizes specific entities within the dataset and attempts to draw some conclusions on how those entities should be labeled or defined. Common classification algorithms are linear classifiers, support vector machines (SVM), decision trees, k-nearest neighbor, and random forest, which are described in more detail below.

- ***Regression*** is used to understand the relationship between dependent and independent variables. It is widely used to make projections, such as for sales revenue for a given business. Linear regression, logistic regression, and polynomial regression are popular regression algorithms.

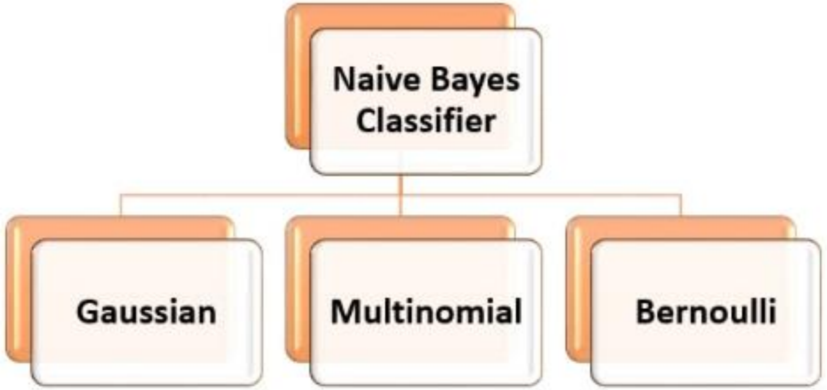
Supervised Learning

Several algorithms and computation techniques are used in supervised machine learning processes. Below are brief explanations of some of the most commonly used learning methods:

1. Naive Bayes
2. Linear regression
3. Logistic regression
4. Support vector machine (SVM)
5. K-nearest neighbor
6. Random forest
7. Neural networks

Supervised Learning - Naive Bayes

Naive Bayes is classification approach that adopts the principle of class conditional independence from the Bayes Theorem. This means that the presence of one feature does not impact the presence of another in the probability of a given outcome, and each predictor has an equal effect on that result. There are three types of Naïve Bayes classifiers: Multinomial Naïve Bayes, Bernoulli Naïve Bayes, and Gaussian Naïve Bayes.



Supervised Learning - Naive Bayes

Multinomial Naive Bayes is used mainly for document classification problem, i.e whether a document belongs to the category of sports, politics, technology etc. The features/predictors used by the classifier are the frequency of the words present in the document.

Gaussian Naive Bayes: When the predictors take up a continuous value and are not discrete, we assume that these values are sampled from a gaussian distribution.

Supervised Learning - Naive Bayes

Bernoulli Naive Bayes only takes binary values. The most general example is where we check if each value will be whether or not a word that appears in a document. That is a very simplified model. In cases where counting the word frequency is less important, Bernoulli may give better results. In simple words, we have to count every value binary term occurrence features i.e. a word occurs in a document or not. These features are used rather than finding the frequency of a word in the document.

Supervised Learning - Linear & Logistic regression

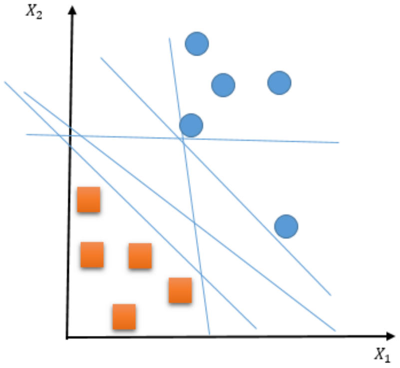
Linear regression is used to identify the relationship between a dependent variable and one or more independent variables and is typically leveraged to make predictions about future outcomes. When there is only one independent variable and one dependent variable, it is known as simple linear regression. As the number of independent variables increases, it is referred to as multiple linear regression. For each type of linear regression, it seeks to plot a line of best fit, which is calculated through the method of least squares. However, unlike other regression models, this line is straight when plotted on a graph.

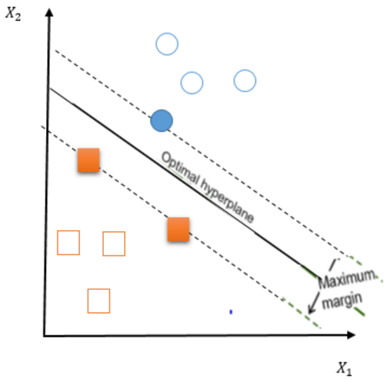
Supervised Learning - Linear & Logistic regression

While linear regression is leveraged when dependent variables are continuous, ***logistic regression*** is selected when the dependent variable is categorical, meaning they have binary outputs, such as "true" and "false" or "yes" and "no." While both regression models seek to understand relationships between data inputs, logistic regression is mainly used to solve binary classification problems, such as spam identification.

Supervised Learning - Support vector machine (SVM)

A support vector machine is a popular supervised learning model developed by Vladimir Vapnik, used for both data classification and regression. That means, it is typically leveraged for classification problems, constructing a hyperplane where the distance between two classes of data points is at its maximum. This hyperplane is known as the decision boundary, separating the classes of data points (e.g., oranges vs. apples) on either side of the plane.





Supervised Learning - K-nearest neighbor

K-nearest neighbor, also known as the KNN algorithm, is a non-parametric algorithm that classifies data points based on their proximity and association to other available data. This algorithm assumes that similar data points can be found near each other. As a result, it seeks to calculate the distance between data points, usually through Euclidean distance, and then it assigns a category based on the most frequent category or average.

Its ease of use and low calculation time make it a preferred algorithm by data scientists, but as the test dataset grows, the processing time lengthens, making it less appealing for classification tasks. KNN is typically used for recommendation engines and image recognition.

Supervised Learning - K-nearest neighbor

Here are some of the advantages of using the k-nearest neighbors algorithm:

- It's easy to understand and simple to implement
- It can be used for both classification and regression problems
- It's ideal for non-linear data since there's no assumption about underlying data
- It can naturally handle multi-class cases
- It can perform well with enough representative data

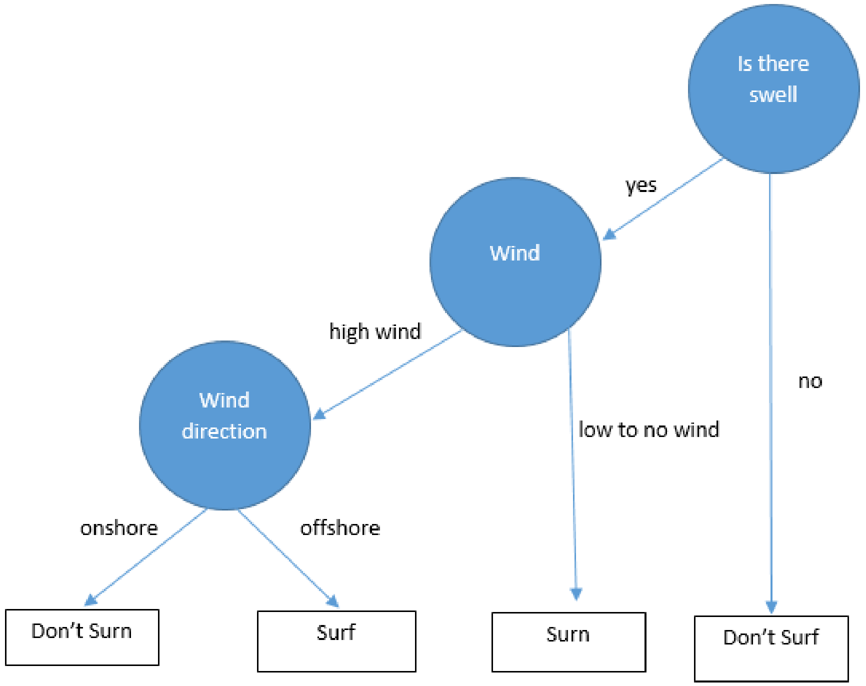
Supervised Learning - K-nearest neighbor

Here are some of the disadvantages of using the k-nearest neighbors algorithm:

- Associated computation cost is high as it stores all the training data
- Requires high memory storage
- Need to determine the value of K
- Prediction is slow if the value of N is high
- Sensitive to irrelevant features

Supervised Learning - Random forest

Random forest is another flexible supervised machine learning algorithm used for both classification and regression purposes. The "forest" references a collection of uncorrelated decision trees, which are then merged together to reduce variance and create more accurate data predictions.



Supervised Learning - Neural Networks

Primarily leveraged for deep learning algorithms, ***neural networks*** process training data by mimicking the interconnectivity of the human brain through layers of nodes. Each node is made up of inputs, weights, a bias (or threshold), and an output. If that output value exceeds a given threshold, it “fires” or activates the node, passing data to the next layer in the network. Neural networks learn this mapping function through supervised learning, adjusting based on the loss function through the process of gradient descent. When the cost function is at or near zero, we can be confident in the model’s accuracy to yield the correct answer. Neural Networks will analyze more in the following slides.

Unsupervised Learning

Machine learning techniques have become a common method to improve a product user experience and to test systems for quality assurance. Unsupervised learning provides an exploratory path to view data, allowing businesses to identify patterns in large volumes of data more quickly when compared to manual observation.

Unsupervised Learning

Some of the most common real-world applications of unsupervised learning are:

News Sections: Google News uses unsupervised learning to categorize articles on the same story from various online news outlets. For example, the results of a presidential election could be categorized under their label for “US” news.

Computer vision: Unsupervised learning algorithms are used for visual perception tasks, such as object recognition.

Unsupervised Learning

Medical imaging: Unsupervised machine learning provides essential features to medical imaging devices, such as image detection, classification and segmentation, used in radiology and pathology to diagnose patients quickly and accurately.

Anomaly detection: Unsupervised learning models can comb through large amounts of data and discover atypical data points within a dataset. These anomalies can raise awareness around faulty equipment, human error, or breaches in security.

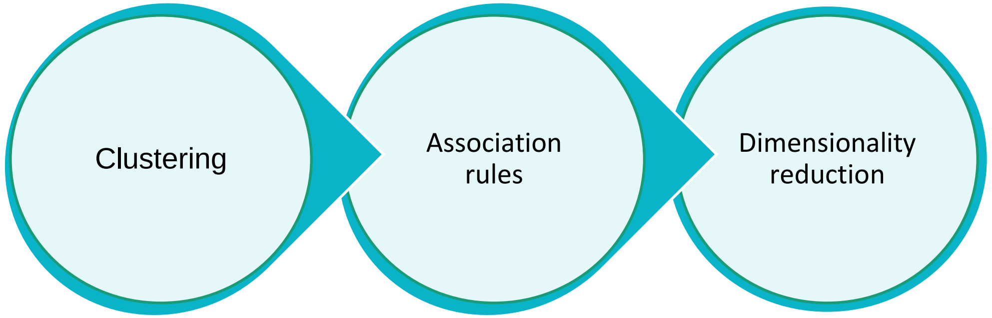
Unsupervised Learning

- ***Customer personas:*** Defining customer personas makes it easier to understand common traits and business clients' purchasing habits. Unsupervised learning allows businesses to build better buyer persona profiles, enabling organizations to align their product messaging more appropriately.

- ***Recommendation Engines***: Using past purchase behavior data, unsupervised learning can help to discover data trends that can be used to develop more effective cross-selling strategies. This is used to make relevant add-on recommendations to customers during the checkout process for online retailers.

Unsupervised Learning

Unsupervised learning, also known as unsupervised machine learning, uses machine learning algorithms to analyze and cluster unlabeled datasets. Its ability to discover similarities and differences in information make it the ideal solution for exploratory data analysis, cross-selling strategies, customer segmentation, and image recognition. Below are presented each learning method and highlight common algorithms and approaches to conduct them effectively.



Unsupervised Learning- Clustering

Clustering is a data mining technique which groups unlabeled data based on their similarities or differences. Clustering algorithms are used to process raw, unclassified data objects into groups represented by structures or patterns in the information. Clustering algorithms can be categorized into a few types, specifically exclusive, overlapping, hierarchical, and probabilistic.

Unsupervised Learning- Clustering

A. Exclusive and Overlapping Clustering

Exclusive clustering is a form of grouping that stipulates a data point can exist only in one cluster. This can also be referred to as “hard” clustering. The K-means clustering algorithm is an example of exclusive clustering. K-means clustering is a common example of an exclusive clustering method where data points are assigned into K groups, where K represents the number of clusters based on the distance from each group’s centroid.

Unsupervised Learning- Clustering

The data points closest to a given centroid will be clustered under the same category. A larger K value will be indicative of smaller groupings with more granularity whereas a smaller K value will have larger groupings and less granularity. K-means clustering is mainly used in market segmentation, document clustering, image segmentation, and image compression.

Overlapping clusters is differentiated from exclusive clustering in that it allows data points to belong to multiple clusters with separate degrees of membership. “Soft” or fuzzy k-means clustering is an example of overlapping clustering.

Unsupervised Learning- Clustering

B. *Hierarchical clustering*

Hierarchical clustering, also known as hierarchical cluster analysis (HCA), is an unsupervised clustering algorithm that can be categorized in two ways; they can be agglomerative or divisive. Agglomerative clustering is considered a “bottoms-up approach.” Its data points are isolated as separate groupings initially, and then they are merged together iteratively on the basis of similarity until one cluster has been achieved.

Unsupervised Learning- Clustering

Four different methods are commonly used to measure similarity:

1. **Ward's linkage:** This method states that the distance between two clusters is defined by the increase in the sum of squared after fusing two clusters into a single cluster.

1. **Average linkage:** This method is determined by the mean distance between two points in each cluster

1. **Complete (or maximum) linkage:** This method is established by the maximum distance between two points in each cluster

1. **Single (or minimum) linkage:** This method is defined by the minimum distance between two points in each cluster

Unsupervised Learning- Clustering

Euclidean distance is the most common metric used to calculate these distances; however, other metrics, such as Manhattan distance, are also cited in clustering literature.

Divisive clustering can be determined as the antipodal of agglomerative clustering; instead it takes a “top-down” approach. On these terms, an individual data cluster is divided based on the differences between data points. Divisive clustering is not frequently used, but it is still worth marking in the context of hierarchical clustering. These clustering methods are usually visualized using a dendrogram, a tree-like diagram that documents the merging or splitting of data points at each iteration.

Unsupervised Learning- Clustering

C. Hierarchical clustering

A probabilistic model is an unsupervised technique that helps us solve density estimation or “soft” clustering problems. In probabilistic clustering, data points are categorized based on the likelihood that they belong to a specific distribution. The Gaussian Mixture Model (GMM) is the one of the most frequently used probabilistic clustering methods. Gaussian Mixture Models are classified as mixture models, which means that they are made up of an unspecified number of probability distribution functions. GMMs are mainly leveraged to define which Gaussian, or normal, probability distribution a given data point belongs to.

Unsupervised Learning- Clustering

If the mean or variance are known, then we can define which distribution a given data point belongs to. Nonetheless, in GMMs, these variables are not known, so we speculate that a latent, or hidden, variable exists to cluster data points accordingly. While it is not required to use the Expectation-Maximization (EM) algorithm, it is a frequently used to estimate the assignment probabilities for a given data point to a specific data cluster.

Unsupervised Learning- Association Rules

An ***association rule*** is a rule-based method for discovering relationships between variables in a given dataset. These methods are commonly used for market basket analysis, allowing companies to better understand relationships among different products. Understanding consumption habits of customers facilitates businesses to develop better cross-selling strategies and recommendation engines. Examples of this can be seen in Amazon's “Customers Who Bought This Item Also Bought” or Spotify's "Discover Weekly" playlist. While there are a few different algorithms used to generate association rules, such as Apriori, Eclat, and FP-Growth, the Apriori algorithm is most widely used.

Unsupervised Learning- Association Rules

A. Apriori algorithms

Apriori algorithms have been popularized through market basket analysis, leading to different recommendation engines for music platforms and online retailers. They are used within transactional datasets to establish frequent itemsets, or collections of items, to identify the likelihood of consuming a product given the consumption of another product. For example, if I play Black Sabbath's radio on Spotify, starting with their song "Orchid", one of the other songs on this channel will expected be a Led Zeppelin song, such as "Over the Hills and Far Away." This is based on my prior listening habits as well as the ones of others.

Unsupervised Learning- Dimensionality reduction

While more data universally yields more accurate results, it can also impact the performance of machine learning algorithms (e.g. overfitting) and it can also make it difficult to visualize datasets. ***Dimensionality reduction*** is a technique used when the number of features, or dimensions, in a provided dataset is too high. It scales down the number of data inputs to a manageable size while also retaining the integrity of the dataset as much as possible.

Unsupervised Learning- Dimensionality reduction

A. Principal component analysis

Principal component analysis (PCA) is a type of dimensionality reduction technique which is used to decrease redundancies and to constrict datasets through feature extraction. This method uses a linear transformation to create a new data representation, yielding a set of "principal components." The first principal component is the direction which maximizes the variance of the dataset. While the second principal component also finds the maximum variance in the data, it is completely uncorrelated to the first principal component, yielding a direction that is perpendicular, or orthogonal, to the first component.

Unsupervised Learning- Dimensionality reduction

B. Singular value decomposition

Singular value decomposition (SVD) is some other dimensionality reduction approach which factorizes a matrix, A, into three, low-rank matrices. SVD is denoted by the formula, A = USVT, where U and V are orthogonal matrices. S is a diagonal matrix, and S values are considered singular values of matrix A. Similar to PCA, it is frequently used to reduce noise and compress data, such as image files.

Unsupervised Learning- Dimensionality reduction

C. **Autoencoders**

Autoencoder is an unsupervised artificial neural network that learns how to efficiently compress and encode data then learns how to reconstruct the data back from the reduced encoded representation to a representation that is as close to the original input as possible.

Autoencoder, by design, reduces data dimensions by learning how to ignore the noise in the data.

Deep Learning-Neural Networks

Models in this area set themselves apart from classical models by increasing model complexity confoundedly. Consolidating and contributing to developments in this area into the TM will speed up the ability to analyze and understand the Big Data of the Past. Specifically, research toward a Universal Representation Space that is able to describe the meaning and semantics of objects, text, images, and other information sources and is able to transfer all of these representations into each other extending on recent progress in machine translation is a major aim TM.

Deep Learning-Neural Networks

The most popular deep learning algorithms are:

- Convolutional Neural Networks (CNNs)
- Long Short Term Memory Networks (LSTMs)
- Recurrent Neural Networks (RNNs)
- Multilayer Perceptrons (MLPs)

Convolutional Neural Networks (CNNs)

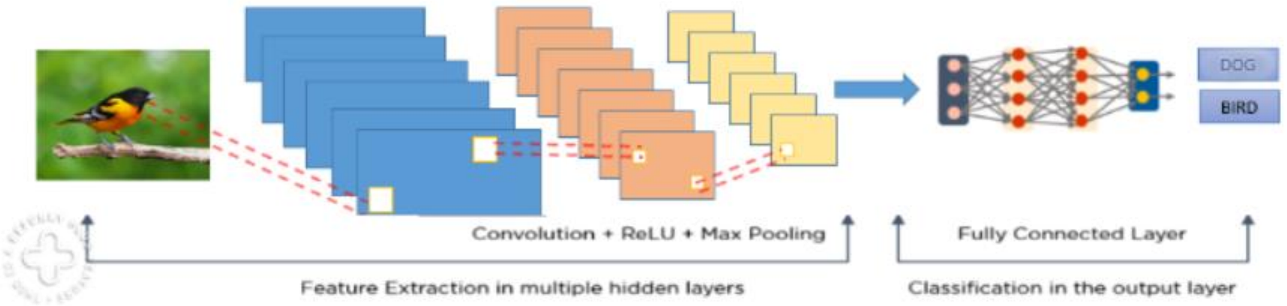
Convolutional Neural Networks or CNNs, also generally known as ConvNets, comprise of multiple layers and are mostly used for image processing and object detection. CNN have multiple layers that process and extract features from data. These are:

- Convolution Layer: CNN has a convolution layer that has several filters to perform the convolution operation.
- Rectified Linear Unit (ReLU): CNN's have a ReLU layer to perform operations on elements. The output is a rectified feature map.

Convolutional Neural Networks (CNNs)

- Pooling Layer: Pooling is a down-sampling operation that lower the dimensions of the feature map. The pooling layer then transform the resulting two-dimensional arrays from the pooled feature map into a single, long, continuous, linear vector by flattening it.

- Fully Connected Layer: A fully connected layer forms when the flattened matrix from the pooling layer is fed as an input, which classifies and identifies the images.



Long Short Term Memory Networks (LSTMs)

LSTMs are a type of Recurrent Neural Network (RNN) that can learn and memorize long-term dependencies. Recalling past information for long periods is the default behavior.

LSTMs preserve information over time. They are useful in time-series prediction because they remember previous inputs. LSTMs have a chain-like structure where four interacting layers communicate in a unique way. Besides time-series predictions, LSTMs are mostly used for speech recognition, music composition, and pharmaceutical development.

Long Short Term Memory Networks (LSTMs)

The LSTM model is a powerful recurrent neural system specifically designed to overcome the exploding/vanishing gradient problems that typically arise when learning long-term dependencies, even when the diminutive time lags are very long. Overall, this can be anticipated by using a constant error carousel (CEC), which maintains the error signal within each unit's cell. As a matter of fact, such cells are recurrent networks themselves, with interesting architecture in the way that the CEC is stretched with additional features, namely the input gate and output gate, forming the memory cell. The self-recurrent connections indicate feedback with a lag of a one-time step.

Long Short Term Memory Networks (LSTMs)

Long Short Term Memory Network compose of four different gates for different objectives as described below:

- Forget Gate: It defines to what extent to forget the previous data.
- Input Gate: It determines the extent of information written onto the Internal Cell State.

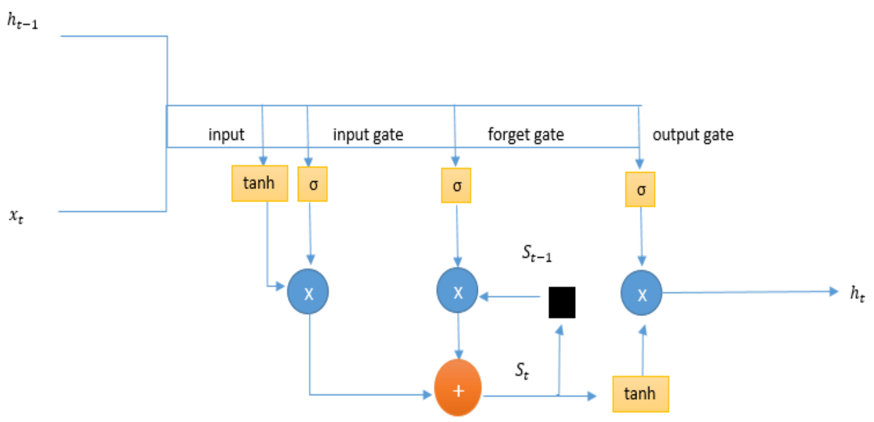
Long Short Term Memory Networks (LSTMs)

- Input Modulation Gate: It is often treated as a sub-part of the input gate and much literature on LSTMs does not even mention it and speculates it is inside the Input gate. It is used to inflect the information that the Input gate will write onto the Internal State Cell by adding non-linearity to the information and making the information Zero-mean. This is operated to decrease the learning time as Zero-mean input has more quick convergence. Although this gate's actions are minor than the others and are often treated as a finesse-providing concept, it is good practice to hold this gate in the structure of the LSTM unit.
- Output Gate: It rules what output(next Hidden State) to create from the current Internal Cell State.

Long Short Term Memory Networks (LSTMs)

The basic workflow of a Long Short Term Memory Network is analogous to the workflow of a Recurrent Neural Network with the only difference being that the Internal Cell State is also passed ahead along with the Hidden State. The LSTMs work as follows:

- First, they forget extraneous parts of the previous state
- Next, they selectively update the cell-state values
- Finally, the output of certain parts of the cell state



Recurrent Neural Networks (RNNs)

RNNs have connections that structure directed cycles, which allow the outputs from the LSTM to be fed as inputs to the current phase.

The output from the LSTM becomes an input to the current phase and can memorize previous inputs due to its internal memory. RNNs are frequently used for image captioning, time-series analysis, natural-language processing, handwriting recognition, and machine translation.

Recurrent Neural Networks (RNNs)

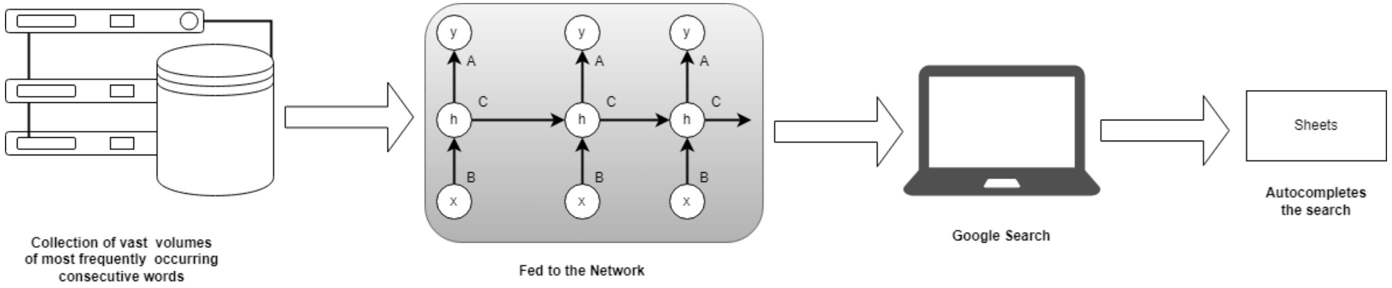
Recurrent neural networks (RNNs) are a superset of feedforward neural networks, expanded with the ability to pass information across time steps. They are a rich family of models adequate for nearly arbitrary computation. In practice, the capability to model temporal dependencies makes recurrent neural networks especially suited to tasks where input and/or output consist of sequences of points that are dependent.

Recurrent Neural Networks (RNNs)

The RNN works as follows:

- The output at time t-1 feeds into the input at time t.
- Similarly, the output at time t feeds into the input at time t+1.
- RNNs can process inputs of any length.
- The computation accounts for historical information, and the model size does not increase with the input size.

Here is an example of how Google's autocompleting feature works:



Multilayer Perceptrons (MLPs)

MLPs are an excellent place to start learning about deep learning technology.

MLPs belong to the class of feedforward neural networks with multiple layers of perceptrons that have activation functions. MLPs consist of an input layer and an output layer that are fully connected. They have the same number of input and output layers but may have multiple hidden layers and can be used to create speech-recognition, image-recognition, and machine-translation software. The input layer acquires the input signal to be processed. The required task such as prediction and classification is performed by the output layer. An willful number of hidden layers that are placed in between the input and output layers are the true computational engine of the MLP.

Multilayer Perceptrons (MLPs)

Alike to a feed-forward network in an MLP the data flows in the forward direction from the input to the output layer. The neurons in the MLP are trained with the back propagation learning algorithm. MLPs are designed to approximate any continuous function and can solve problems that are not linearly fissile.

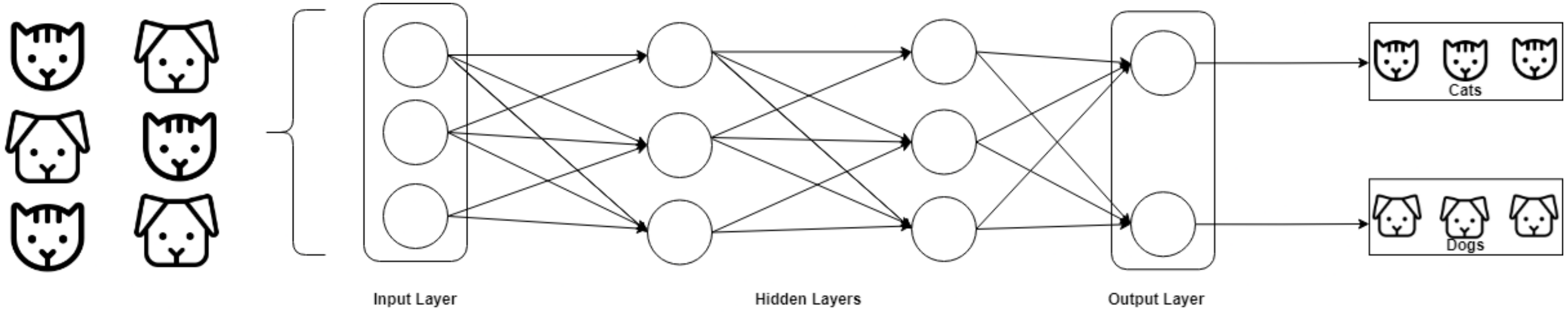
Multilayer Perceptrons (MLPs)

The MLPs work as follows:

- MLPs feed the data to the input layer of the network. The layers of neurons associate in a graph so that the signal passes in one direction.
- MLPs calculate the input with the weights that exist between the input layer and the hidden layers.

Multilayer Perceptrons (MLPs)

- MLPs use activation functions to define which nodes to fire. Activation functions comprise ReLUs, sigmoid functions, and tanh.
- MLPs train the model to understand the correlation and know about the dependencies between the independent and the target variables from a training data set.



Other prominent AI techniques - Reinforcement Learning

Reinforcement Learning deals with learning via interaction and feedback, or in other words learning to solve a task by trial and error, or in other-other words acting in an environment and receiving rewards for it. Originally, an agent (or several) is built that can perceive and interpret the environment in which is placed, furthermore, it can take actions and interact with it.

Other prominent AI techniques - Reinforcement Learning

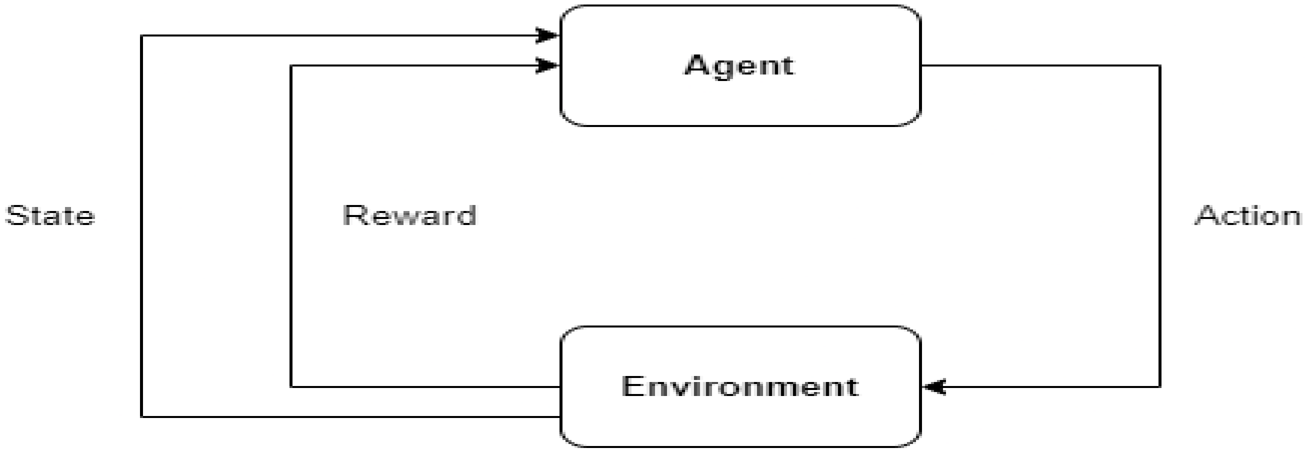
- **Agent:** The learner and the decision maker.
- **Environment:** Where the agent learns and establishes what actions to perform.
- **Action:** A set of actions which the agent can execute.
- **State:** The state of the agent in the environment.
- **Reward:** For each action selected by the agent the environment provides a reward. Usually a scalar value.

Other prominent AI techniques - Reinforcement Learning

- **Policy (π):** It is a strategy which implemented by the agent to decide the next action based on the current state.
- **Value (V):** It is expected long-term return with discount, as compared to the short-term reward.
- **Value Function:** It designates the value of a state that is the total amount of reward. It is an agent which should be expected beginning from that state.
- **Model of the environment:** This mimics the behavior of the environment. It helps you to make inferences to be made and also determine how the environment will behave.

Other prominent AI techniques - Reinforcement Learning

- **Model based methods:** It is a method for solving reinforcement learning problems which use model-based methods.
- **Q value or action value (Q):** Q value is quite similar to value. The only difference between the two is that it takes an additional parameter as a current action.



Reinforcement Learning process

In a way, Reinforcement Learning is the science of making optimal decisions using experiences. Breaking it down, the process of Reinforcement Learning associate the following simple steps:

- Observation of the environment
- Deciding how to act using some strategy
- Acting appropriately
- Receiving a reward or penalty
- Learning from the experiences and refining our strategy
- Iterate until an optimal strategy is found

AI Applications of Today

The AI used in many sectors by helping for better decision making, increasing performance and eliminating repetitive work. The main sectors that AI is commonly used are:

- **Healthcare:** In Healthcare devices like smart watches which records sleep patterns, vital signs uses AI in order to monitor and notify abnormal situations.

- **Automobile:** The self driving is growing with the help of AI. An example of this is the autopilot by Tesla that takes up data from all the Tesla's running on the road and uses it in machine learning algorithms.

AI Applications of Today

- **Banking and Finance:** Features like AI bots, digital payment advisers and biometric fraud detection mechanisms cause higher quality of services to a wider customer base. The adoption of AI in banking is constant to rework companies within the industry, provide greater levels useful and more personalized experiences to their customers, reduce risks.

AI Applications of Today

- **Surveillance:** AI empowers the systems to monitor the footage in real-time and can be a pathbreaking development in regards to public safety.

- **Social Media:** In social media AI tools work silently within the background, showing us posts that we “might” like and advertising products that “might” be useful based on our search and browsing history.

AI Applications of Today

- **Entertainment:** Online streaming services like Netflix and Amazon Prime, relies heavily on the info collected by the users. This helps with recommendations based upon the previously viewed content. This is done not only to deliver accurate suggestions but also to create content that would be liked by a majority of the viewers.

- **Education:** AI systems easily adapt to each student’s individual learning needs and can target instruction based on their strengths and weaknesses. Also AI helping administrators to schedule courses and individuals to manage their daily, weekly, monthly or yearly schedules.

Examples of AI models by KONNEKTABLE

1. **Long term water consumption:** it is used a state-of-the-art machine learning (Deep Neural Networks) techniques on the problem of time series prediction. It focused on the development of novel approaches that combine high accuracy and interpretable insights in terms of forecasting.

1. **mtune:** This application is used for the optimization of processes in production systems taking into account several parameters needed using artificial intelligence methods and making use of machine learning models.

Examples of AI models by KONNEKTABLE

3. **KYKLOS's DSS:** This DSS has created during the KYKLOS project. DSS is a component with the aim of providing alerts and recommendations on the Short-Term Analysis, which provides alerts coming from the end-users' assets through an alert box. Also the Long-Tern Analysis has been designed to calculate a total circularity potential score for at least two concepts, in order to be compared in terms of circularity. (https://kyklos40project.eu/about-kyklos/technology/)

Examples of AI models by KONNEKTABLE

4. **SOT DSS:** SOT DSS has created during the SnR project. The SOT DSS assists and helps the coordination of the response activities, using data from multiple, heterogeneous sources (historical and real-time data) and Optimization techniques. It is intended to support the end-users regarding decision-making. (https://search-and-rescue.eu/concorde-platform/).

5. **Forest fire DSS**: By using Machine learning algorithms the DSS predict how many firefighters, fire trucks should be sent to a forest fire. Also it predicts the burnt area based on historical data.

Examples of AI models by KONNEKTABLE

4. **SPHINX DSS:** This DSS has developed during the SPHINX project. The DSS has two functionalities:
    - **The Proactive functionality** uses Machine Learning algorithms to predict possible cyber attacks based on network traffic.
    - **The Active functionality** is a rule based system provides courses of action for the detected cyber attacks by using a IF-THEN rules.

Examples of AI models by KONNEKTABLE

6. **Twitter Sentiment Analysis**: It provides access to the Twitter API handling the authorization. It fetches Tweets based on user's keyword. It cleans the text from Tweets by removing links, stopwords and special characters. It then follows the regular NLP procedures for text handling (such as tokenization) and feeds the text to a Naïve Bayes classifier that gets the sentiment through training the model with TextBlob's Corpora. Finally, the current specific implementation prints a sample of Positive and Negative tweets (the whole unprocessed original text) and plots a pie chart with the relevant distribution of percentages.

Bibliography

1. *Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow, 2nd Edition (Aurelien Géron)*
2. *Deep Learning with Python, 2nd Edition (François Chollet)*
3. *Understanding Machine Learning: From Theory to Algorithms (Shai Shalev-Shwartz and Shai Ben-David)*
4. *Reinforcement Learning: An Introduction (Richard S. Sutton and Andrew G. Barto)*
5. *K-Nearest Neighbors Algorithm & Unsupervised Learning, IBM*
6. *David Loshin, in Business Intelligence (Second Edition), 2013*

AI.4.educators

AI.4.educators
Educating Educators on Artificial Intelligence (AI) – development of an AI training material and an AI educational program for educators

Project No: 2021-1-EL01-KA210-ADU-000034976

AI Practical Roadmap

Co-funded by
the European Union