Introduction to AI within Watershed Management

Isabella Bertani, Qian Zhang, Mike Evans

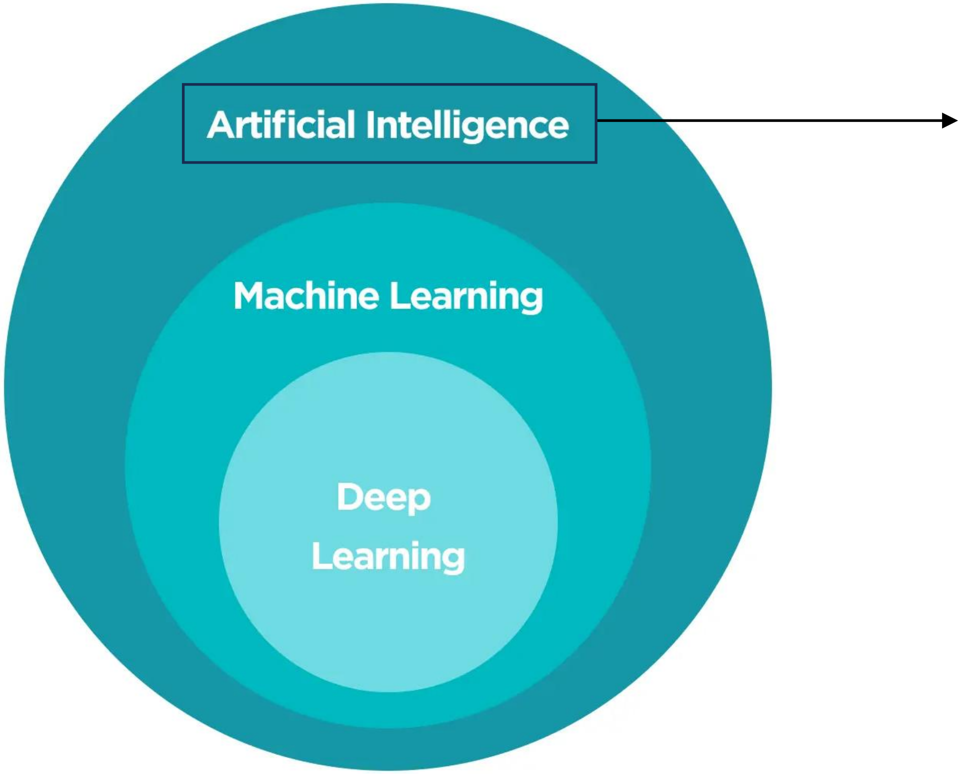
STAC Quarterly Meeting

09/12/2023

Outline

- Isabella Bertani — Broad introduction to Artificial Intelligence, Machine Learning, and Deep Learning
- Qian Zhang — Examples of Machine Learning applications in the Chesapeake Bay Watershed
- Mike Evans — Using Deep Learning to map wetlands and solar arrays in the Chesapeake Bay Watershed

Artificial Intelligence

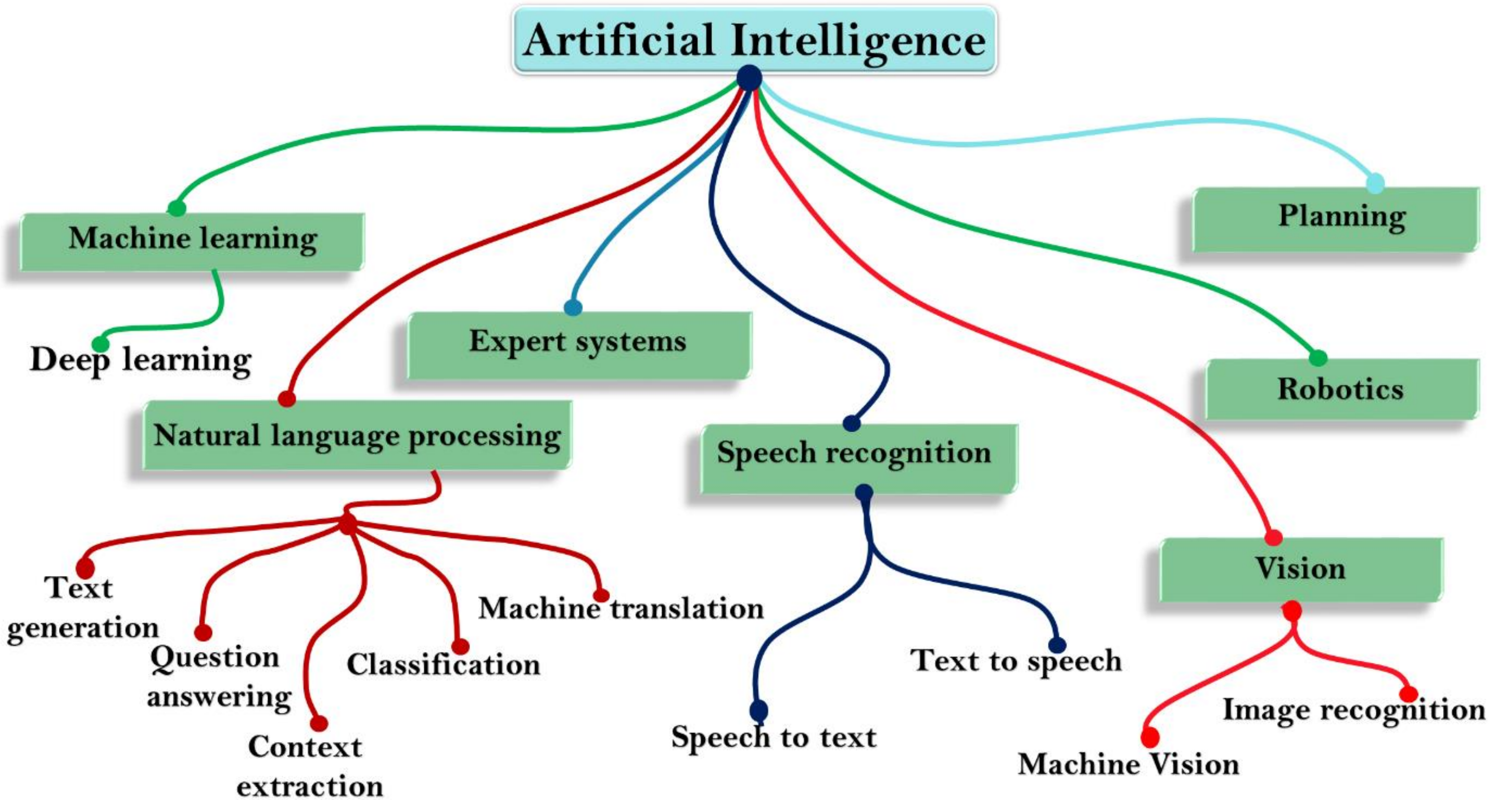


Set of techniques focused on building computers and machines that can reason, learn, and act in ways typically associated with human intelligence

Enables machines to perform cognitive functions such as seeing, understanding, translating and responding to spoken and written language, analyzing data, making recommendations, etc.

Artificial Intelligence

AI is a broad term that encompasses several subfields



https://www.javatpoint.com/subsets-of-ai

Artificial Intelligence

AI is a broad term that encompasses several subfields

<u>Natural Language Processing</u>: focuses on enabling computers to understand, interpret, and generate human language.

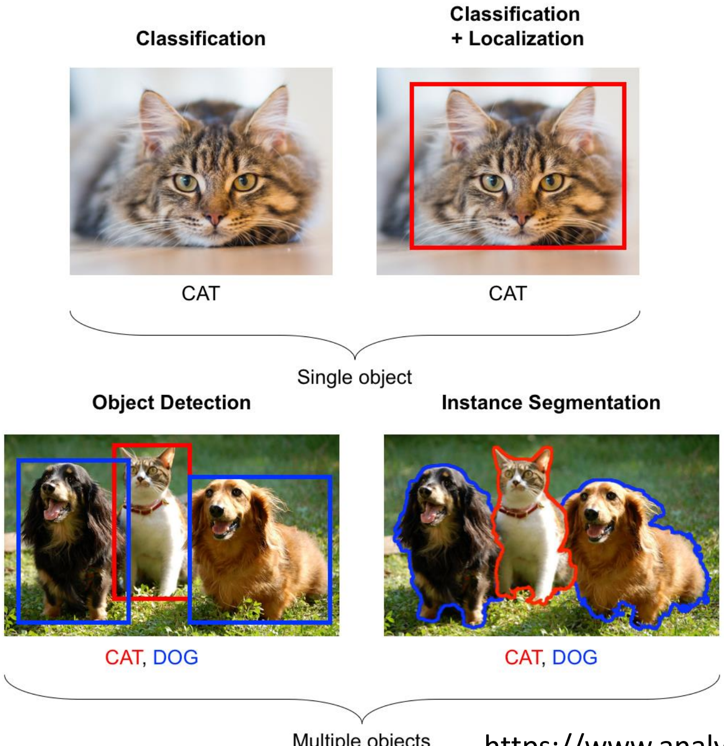
<u>Speech recognition</u>: focuses on enabling machines to understand spoken language and convert it into text

https://datasciencedojo.com/blog/natural-language-processing-applications/

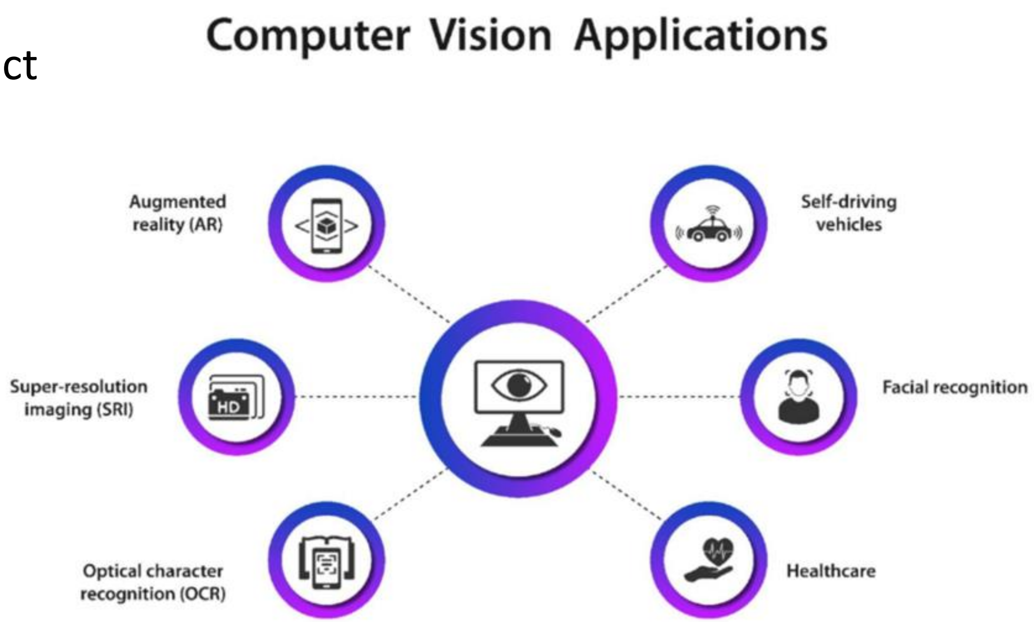
Artificial Intelligence

AI is a broad term that encompasses several subfields

<u>Computer Vision</u>: focuses on enabling computers to process and interpret visual information (e.g., images and videos). Example tasks: image classification, object detection, image segmentation, optical character recognition, etc.

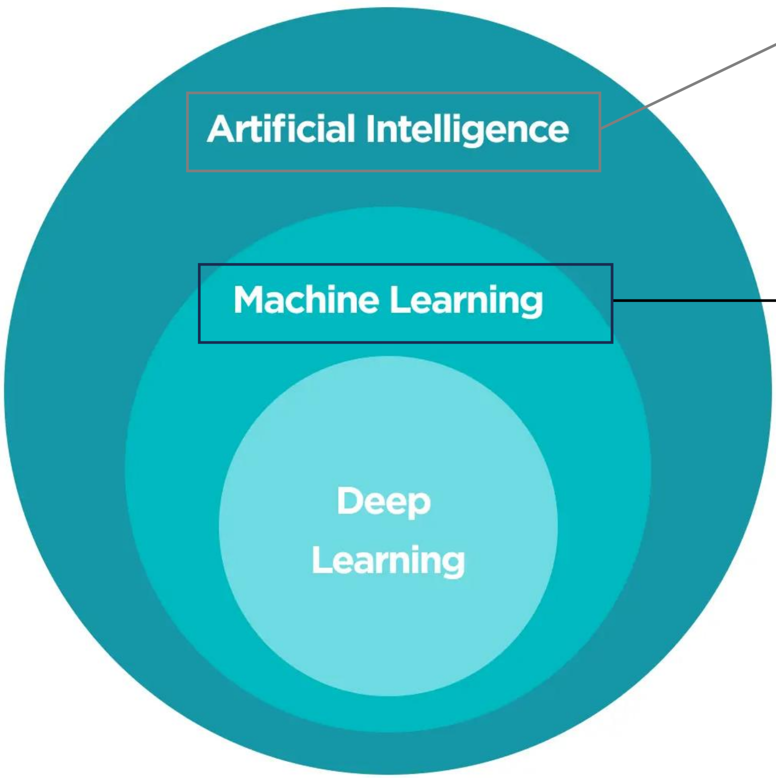


https://www.analyticsvidhya.com/



https://www.nomidl.com/computer-vision/

Machine Learning



Set of techniques focused on building computers and machines that can reason, learn, and act in ways typically associated with human intelligence

A subfield of AI that uses algorithms to enable machines to learn from experience (i.e., data) and make predictions or decisions without being explicitly programmed.

A computer program "learns" from experience with respect to some task if the performance on that task improves with experience.

Machine Learning

Mainly involves training models on data to learn from the data and improve the ability to perform a task. Often the task is to make a prediction or to uncover hidden patterns in data.

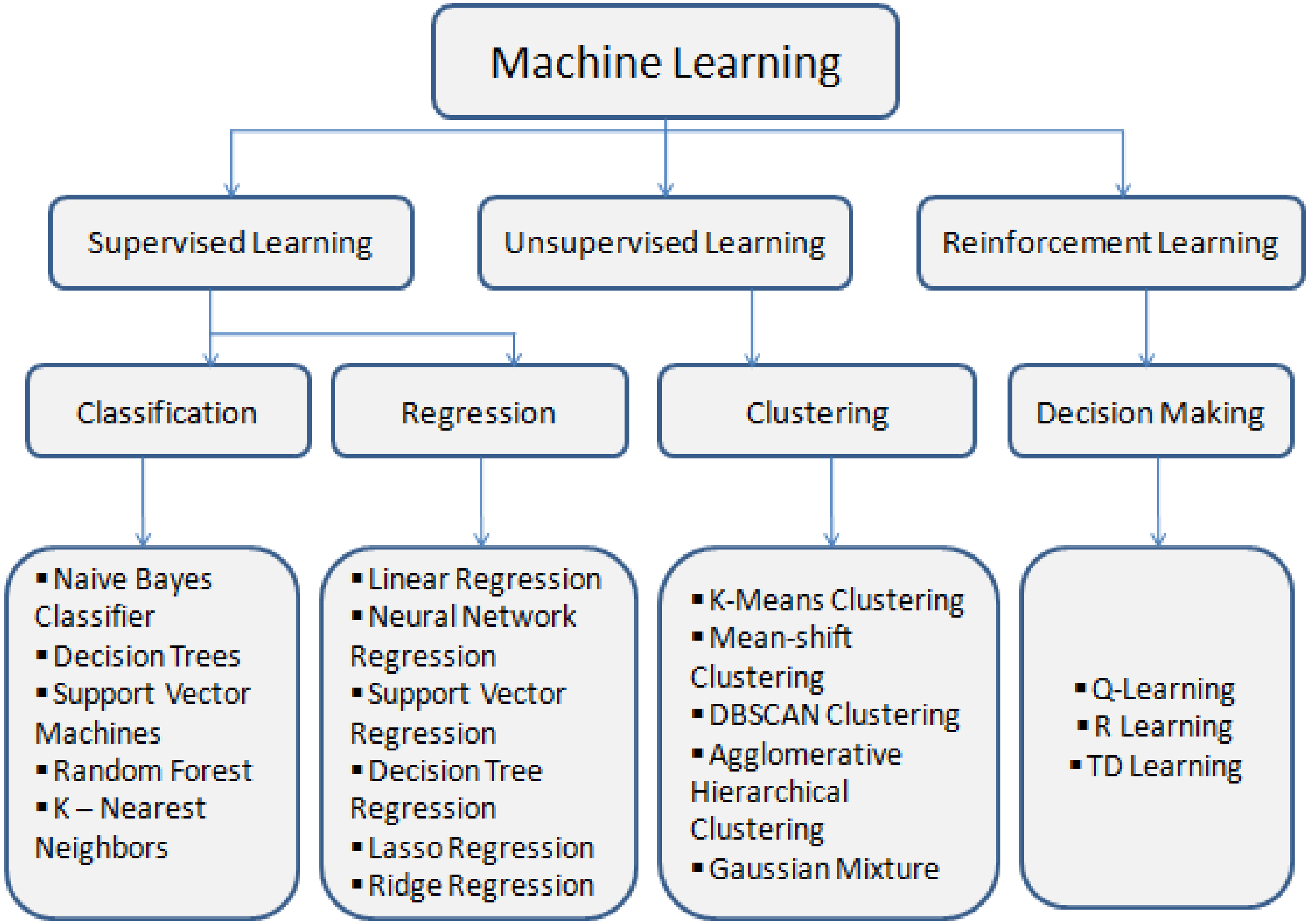
Supervised Learning: algorithms are trained using «labeled data» (i.e., data is tagged with the correct answer). Example: training an e-mail spam detector with a dataset of emails that are labeled as «spam» or «not spam».

Unsupervised Learning: algorithms are trained to discover patterns or form groupings on data that has no labels. Example: group news articles based on their content.

Reinforcement learning: algorithms are trained based on feedback received after each action/decision taken (trial and error). Example: learn to play cards by winning/losing

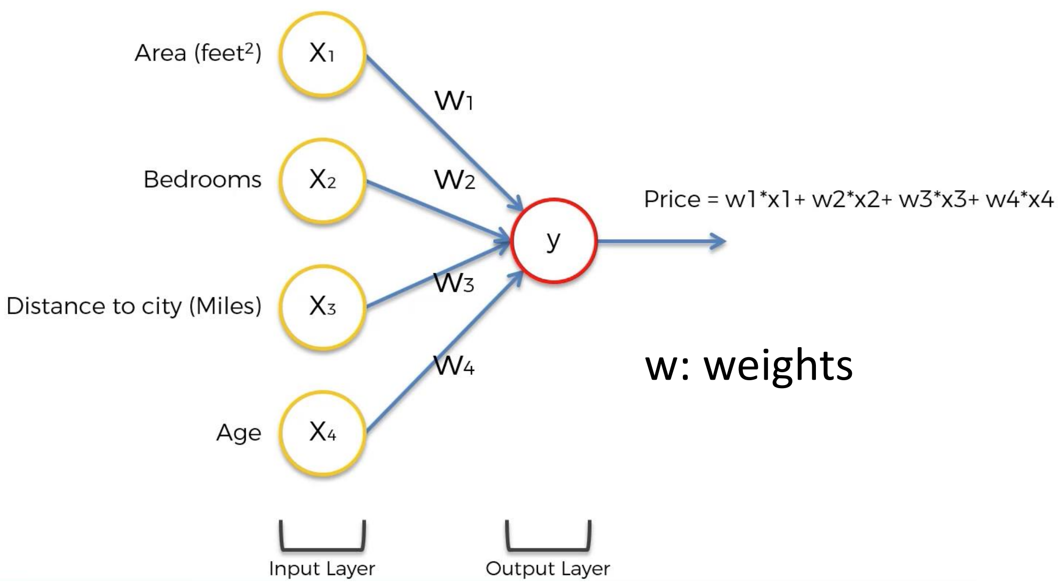
Machine Learning

Common machine learning algorithms



Machine Learning

Example of machine learning algorithm: artificial neural networks



w: weights

Model that mimics the learning behavior of the human brain

Network of interconnected computational units (**neurons)** organized in **layers**.

Input layer: Feeds input variables (predictors) into the model

Output layer: Contains predicted values of the target variable

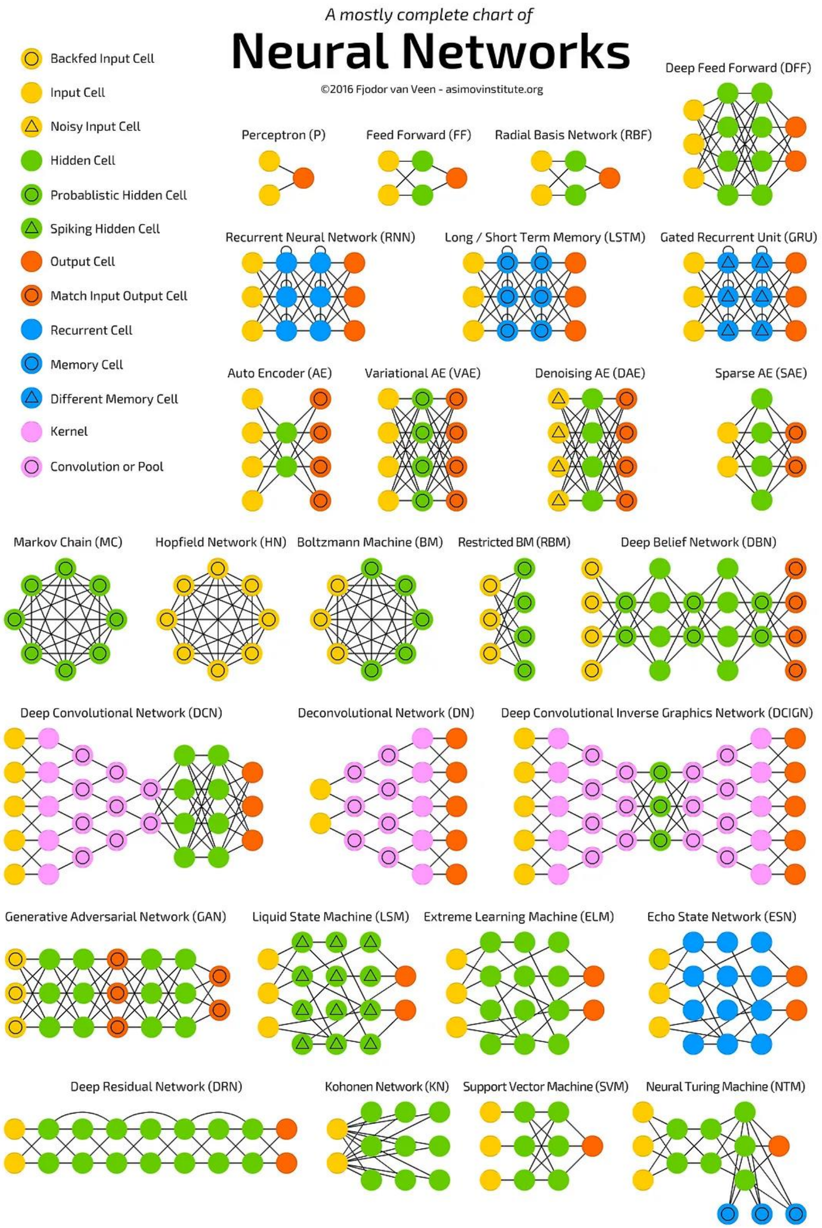
Neurons in a downstream layer take inputs from upstream layer, perform *some computation*, and pass results to downstream layer

Weights are estimated during model training and quantify the importance of each input in predicting the final output

https://www.freecodecamp.org/

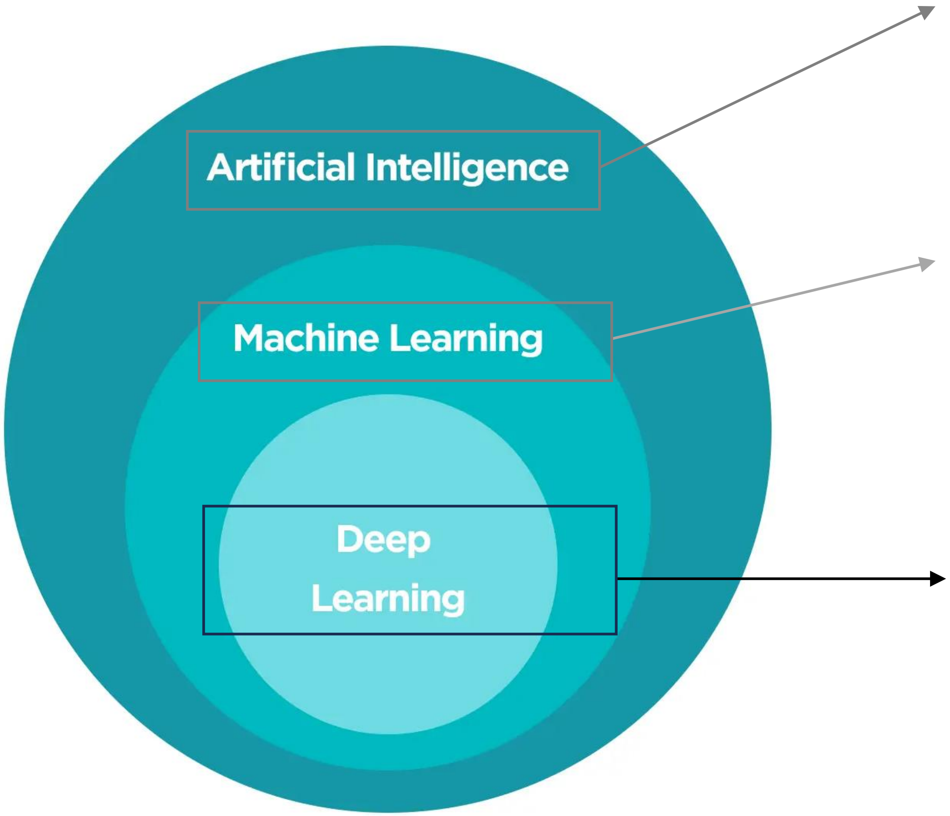
Machine Learning

"The zoo of neural networks"



https://towardsdatascience.com/the-mostly-complete-chart-of-neural-networks-explained-3fb6f2367464

Deep Learning



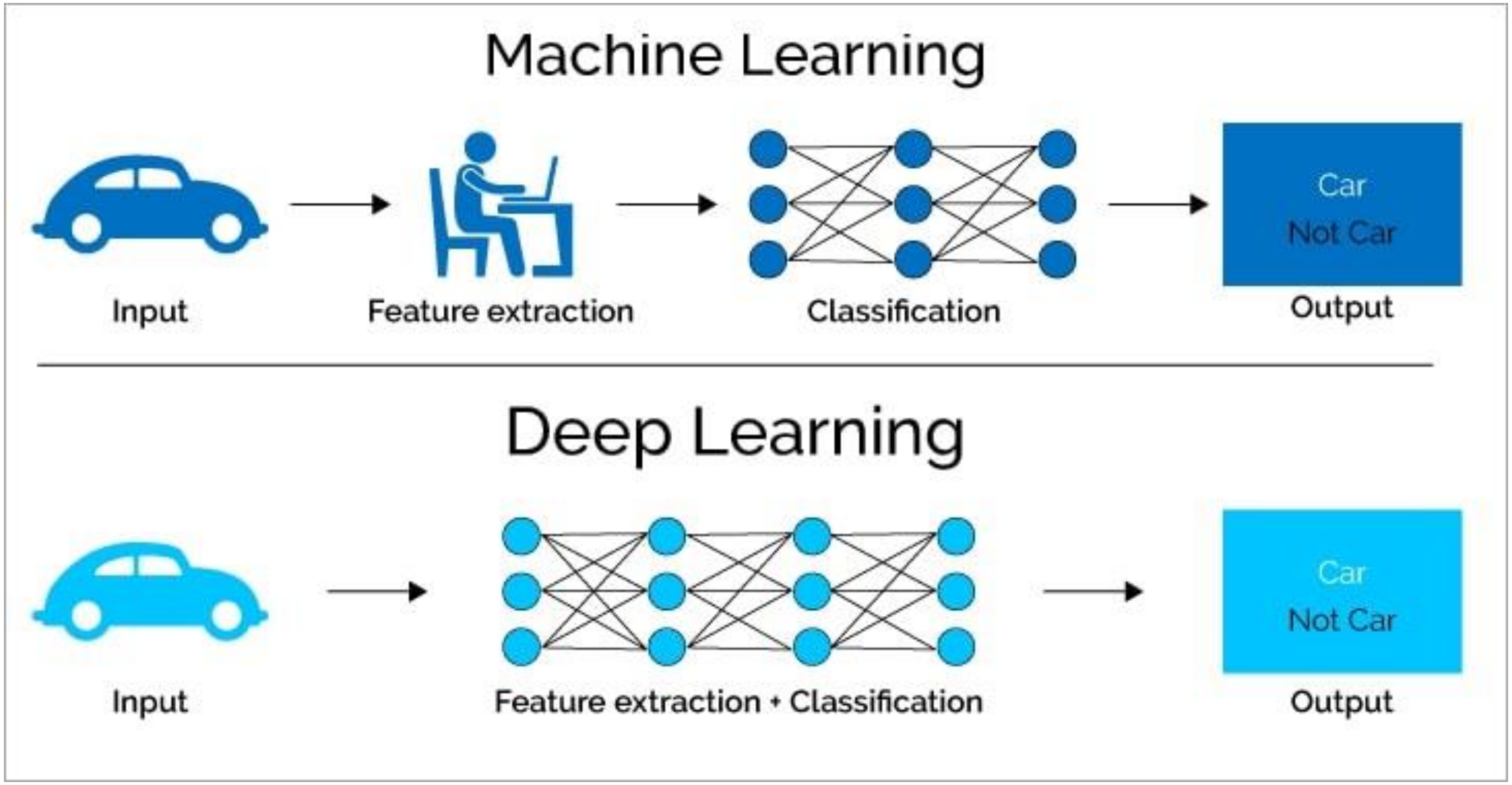
Set of techniques focused on building computers and machines that can reason, learn, and act in ways typically associated with human intelligence

A subfield of AI that uses algorithms to enable machines to learn from experience (i.e., data) and make predictions or decisions without being explicitly programmed

Subset of machine learning algorithms based on advanced artificial neural networks typically characterized by several layers (hence “deep”)

Deep Learning

Compared to more traditional ML algorithms, where humans typically identify what features are important for prediction, DL can process raw, unstructured data (e.g., images), and it automates feature extraction, reducing the need for human pre-processing.



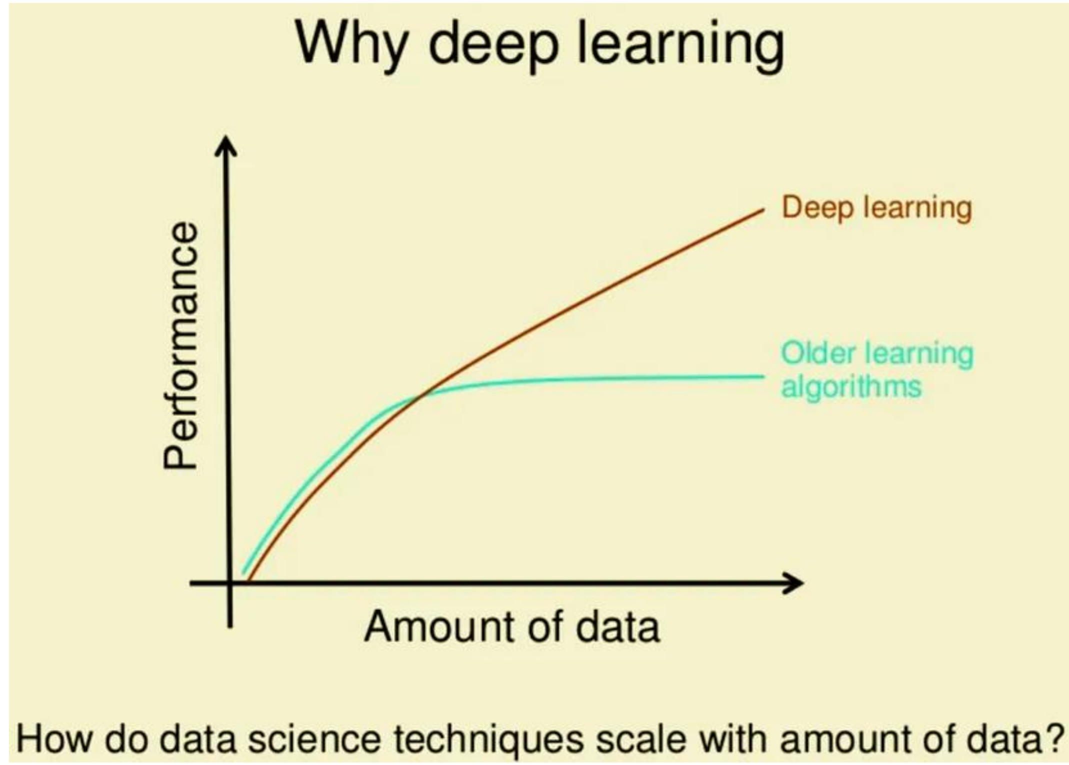
https://levity.ai/blog/difference-machine-learning-deep-learning

Automatic feature engineering has the potential to allow algorithms to discover relationships and concepts beyond what may be conceived by humans.

Deep Learning

DL typically achieves unprecedented predictive accuracy when analyzing large, complex, highly-dimensional datasets, where more traditional statistical methods often fall short

Although more computationally expensive, processing power is evolving rapidly and becoming increasingly available



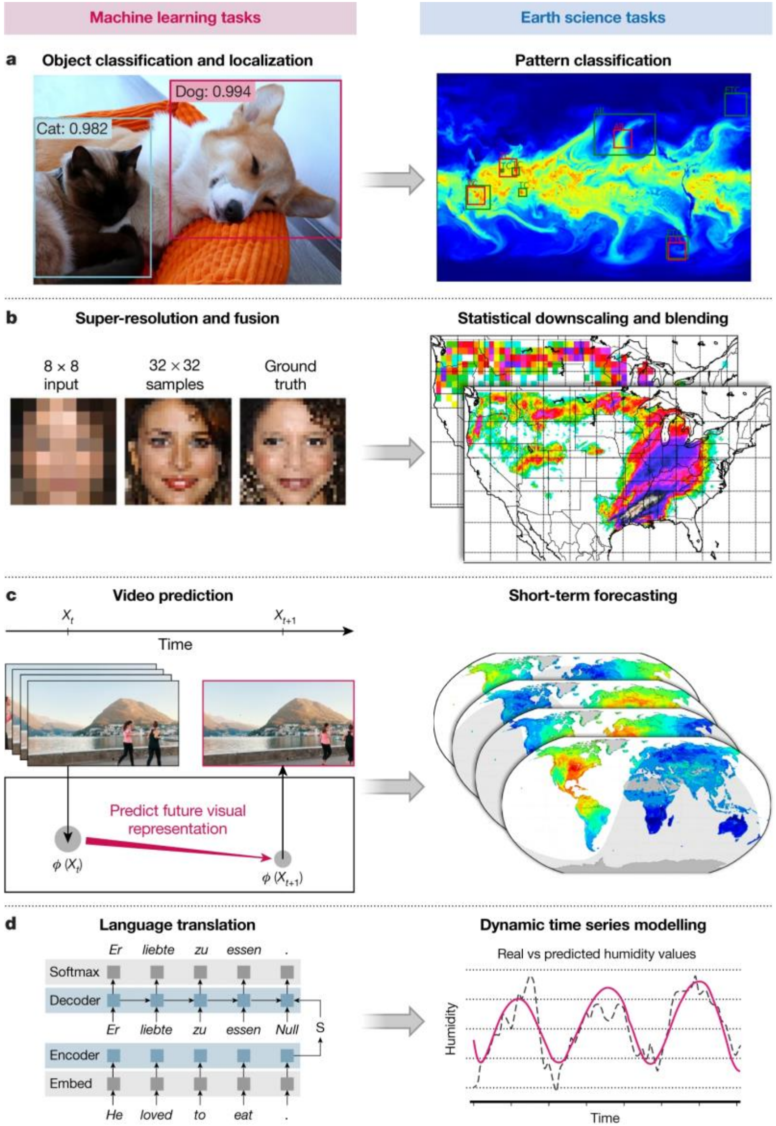
https://towardsdatascience.com

Deep Learning

What is deep learning particularly good at?

Four examples of typical DL applications (left panels) and the geoscientific problems they can be applied to (right panels) (Reichstein et al., 2019)

DL is especially good at information extraction from image-like data (e.g., image classification, object recognition) and sequential data (e.g., speech recognition, time series modeling)



Artificial Intelligence

Advantages

- Solve real-world complex problems easily/fast and efficiently
- Automate (repetitive) tasks and workflows
- Reduce human error/bias and increase accuracy/precision

Concerns

- Data availability and quality
- Sensitive data handling and storing
- Lack of transparency/accountability?