

Data Science

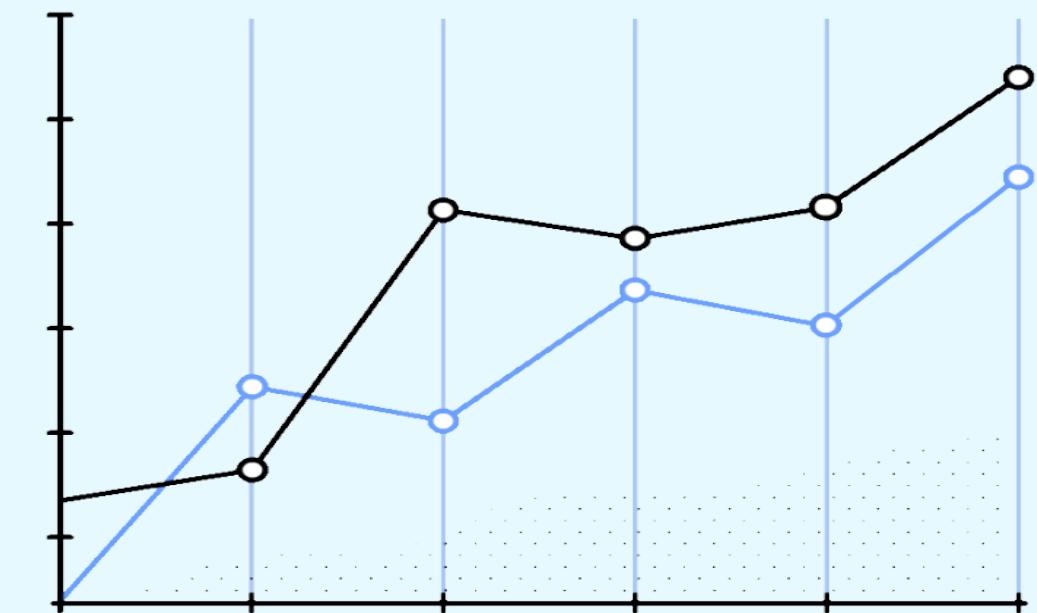


Statistic

Probability

Linear Algebra

Calculus





Maths ?

Descriptive Statistics

Measure of central tendency: Mean, Median, Mode

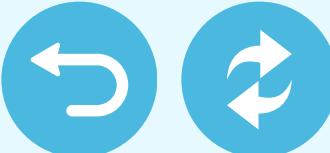
Measure of spread: Range, Standard Deviation, variables, interQuartile Range

Measure of shape: Skewness and kurtosis

Statistical inference: drawing inference about population from sample

Parameter Estimation: Point estimation and confidence interval

Hypothesis testing: One sample hypothesis testing



Maths ?

Differential Statistics

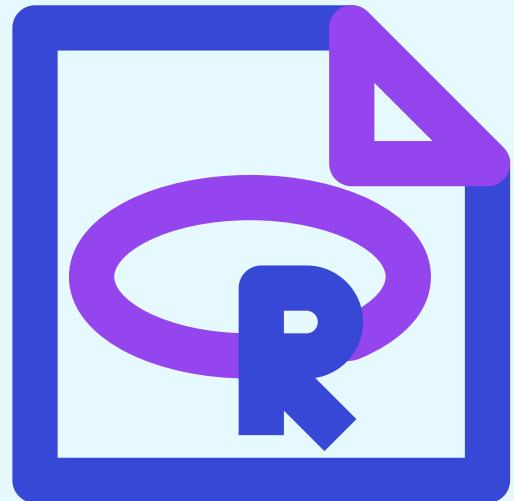
Two sample Hypothesis Testing: ANOVA, MANOVA,
ANCOVSA and MANCOVA

Assosiative Statistics

Finding relationship between two variables,corelation: Pearson, Spearman and Kendall.



Best programming Language

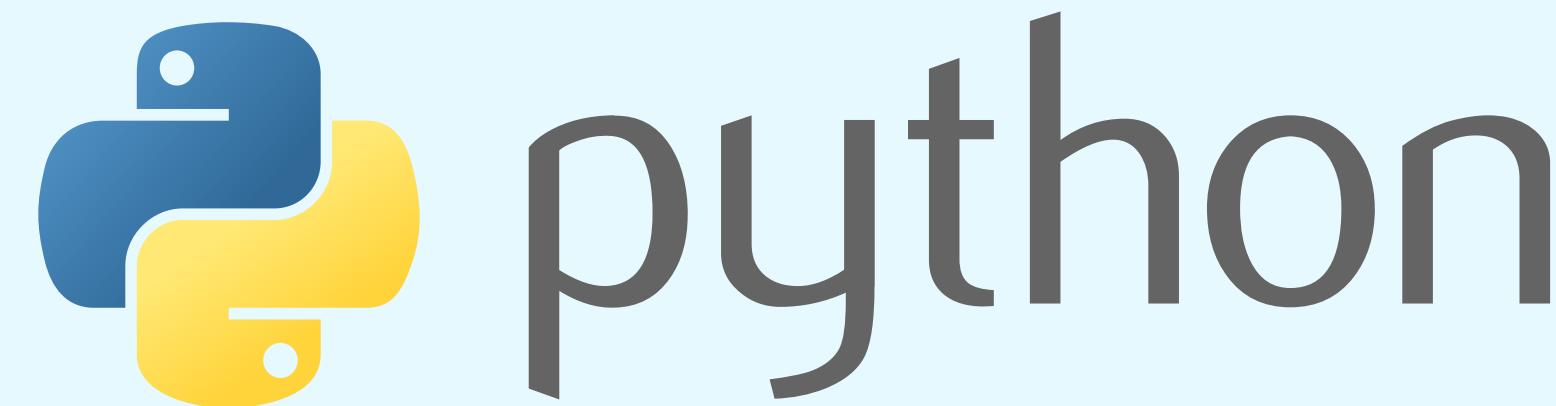


R Programming

- Data Science and Analysis
- Statistical Programming
- Mostly use for Machine learning

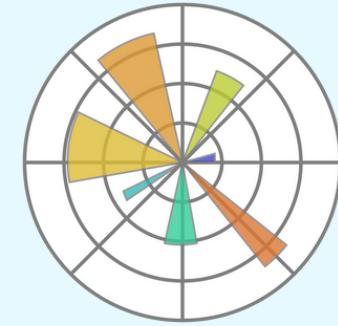
Python Programming

- Use for all purpose and fields
- Superb Libraries
- Ultimate Features
- Communities and support





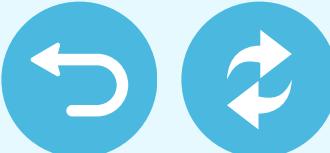
python



Libraries

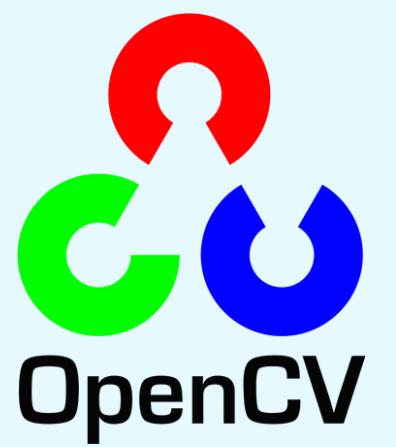
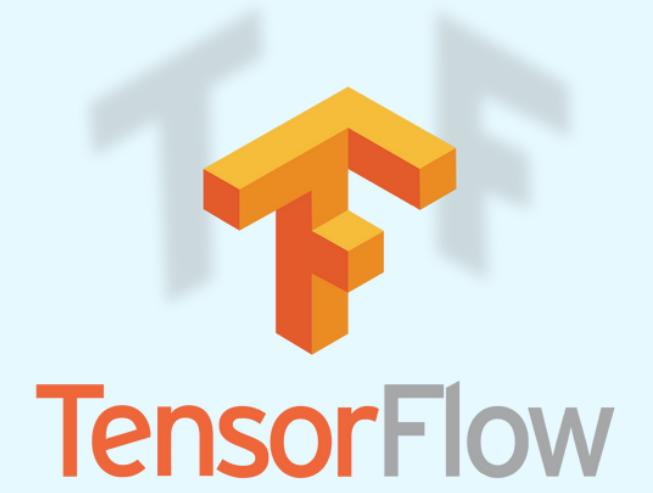
There Are various libraries in python program but we are using some important libraries

- NumPy: use for mathematics terms Ex: Matrices ,Linear Algebra
- Pandas: Data Analysis, Preprocessing,
- Matplotlib: Visualization, Insights
- SeaBorn: Visualization, Insights
- Scikit learn: Build machine Learning Algorithm



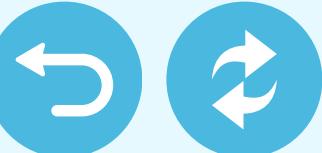
Machine Learning Libraries

- TensorFlow
- Keras
- Nltk
- OpenCv
- PyTorch



NLTK

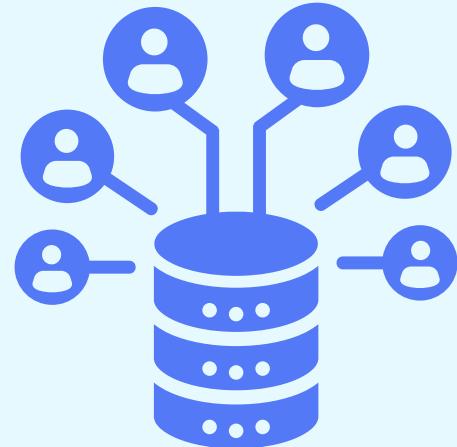




Software Tools

use for view Data analytic with various chart/Graph

- Tableau
- RapidMiner
- KNIME
- Power BI
- QlikView
- Talend
- Splunk
- Apache Spark



Collection Big Data /Demo Data

- Kaggle
- Uci Machine Learning repository
- Google data Set search
- Awesome Public Data Set
- DataWorld
- Data is Plural
- make over monday
- users to find and publish data sets
- explore and build models in a web-based data-science environment
- work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges.

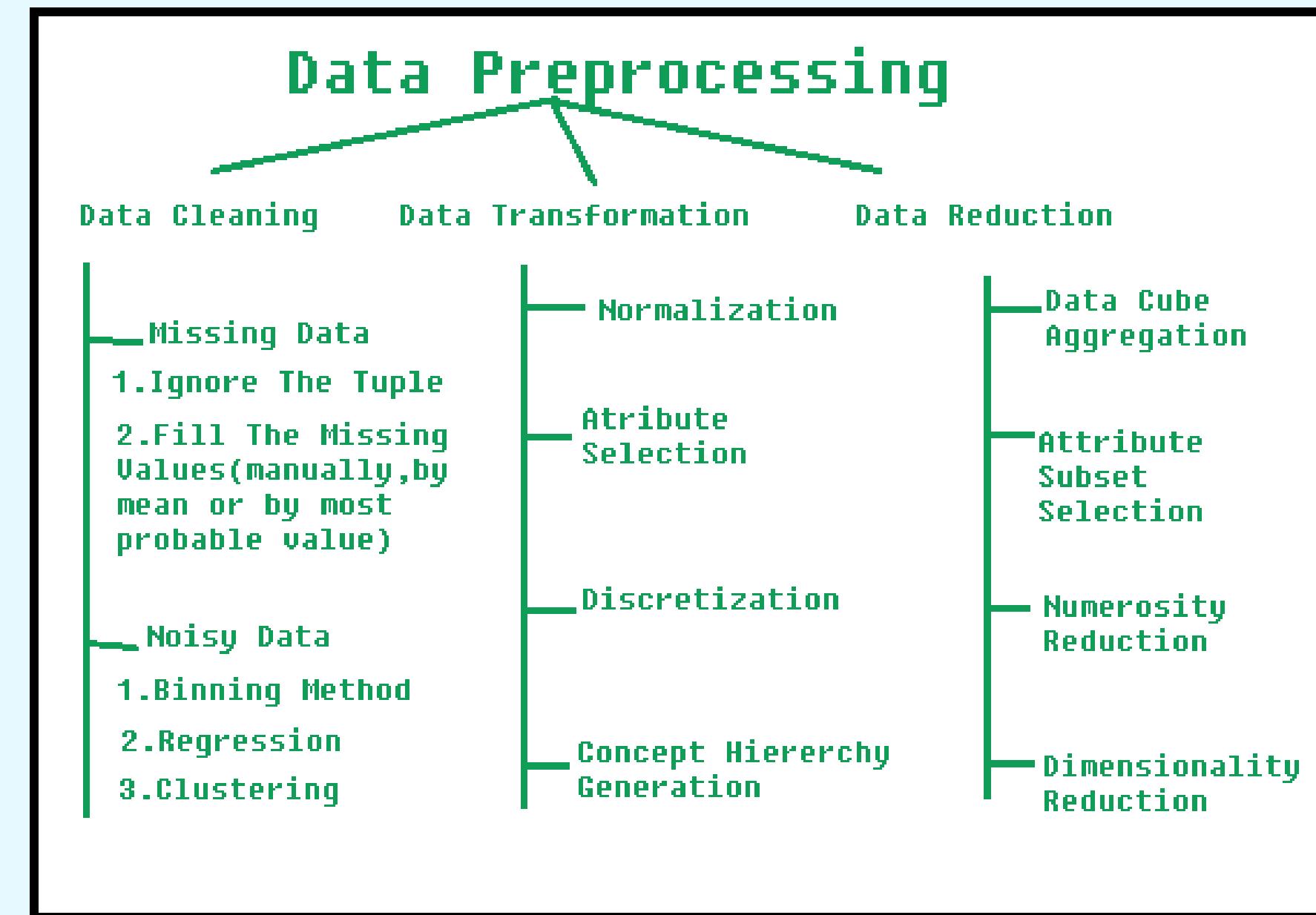


Data PreProcessing

Data preprocessing is a data mining technique which is used to transform the raw data in a useful and efficient format.

Data Formats

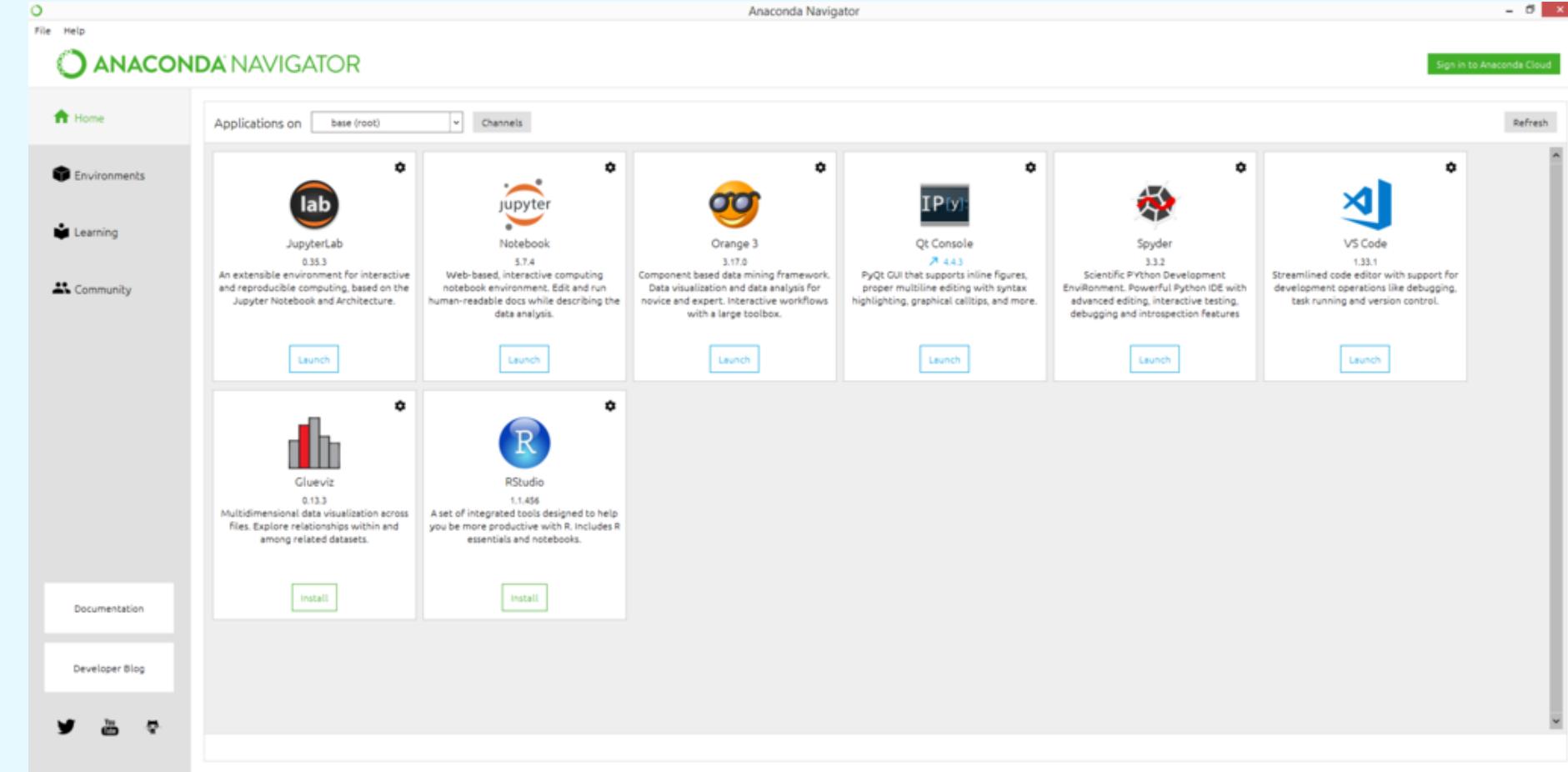
- text
- numbers
- images
- videos
- audio





Best IDE for Work

- Pycharm
- Jupyter Notebook
- Spyder
- Visual Studio Code
- Anaconda



Install Anaconda

Navigator is a desktop graphical user interface (GUI) included in Anaconda distribution that allows users to launch applications and manage conda packages, environments and channels without using command-line commands.



Feature Engineering

Feature engineering is the process of selecting, manipulating, and transforming raw data into features that can be used in supervised learning. In order to make machine learning work well on new tasks, it might be necessary to design and train better features.

Feature engineering, in simple terms, is the act of converting raw observations into desired features using statistical or machine learning approaches.

Feature Engineering Techniques for Machine Learning

1.Imputation

- Numerical Imputation
- Categorical Imputation

2.Handling Outliers

- Removal
- Replacing values
- Capping
- Discretization

3.Log Transform

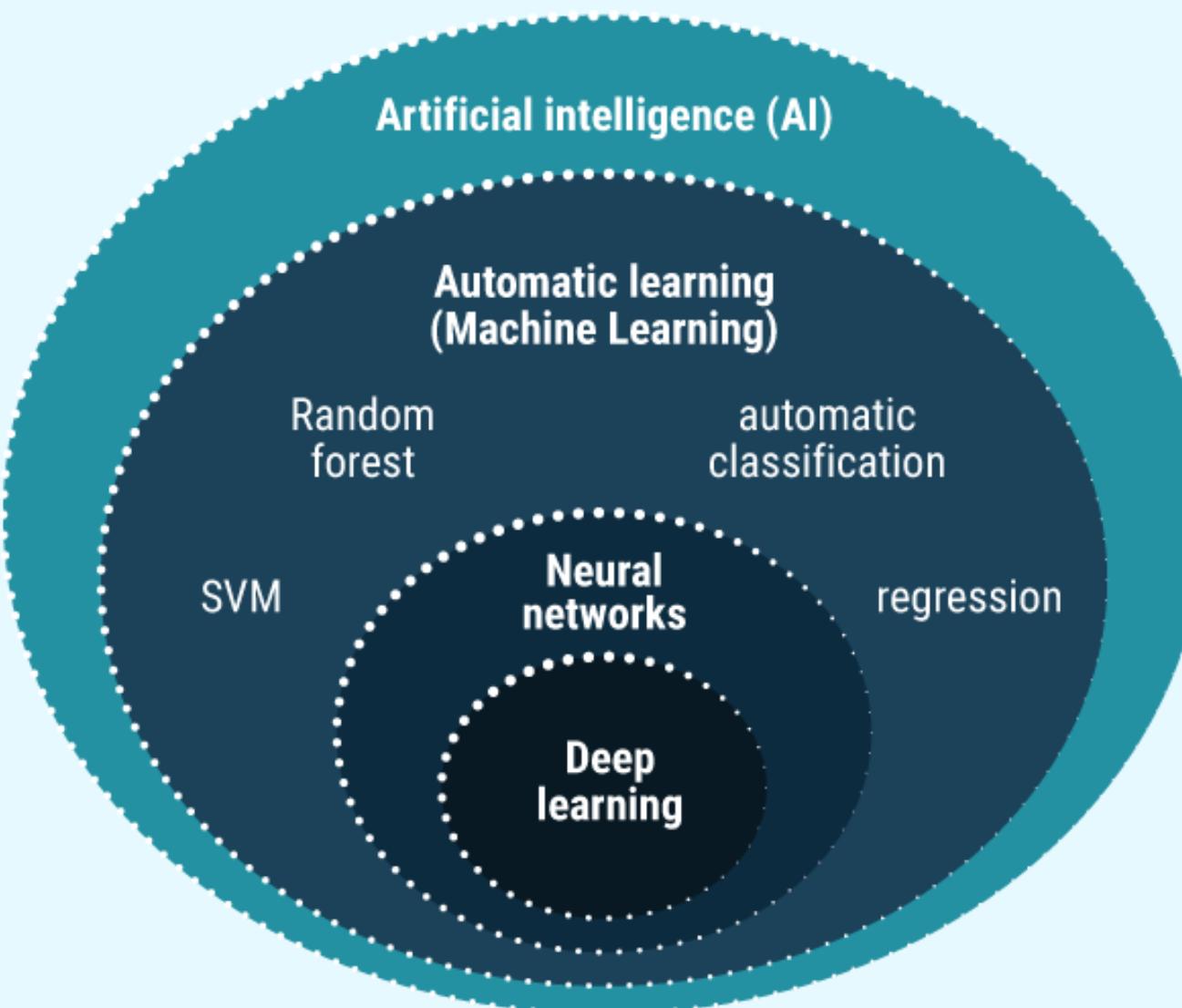
4.One-hot encoding

5.Scaling

- Normalization :
- Standardization

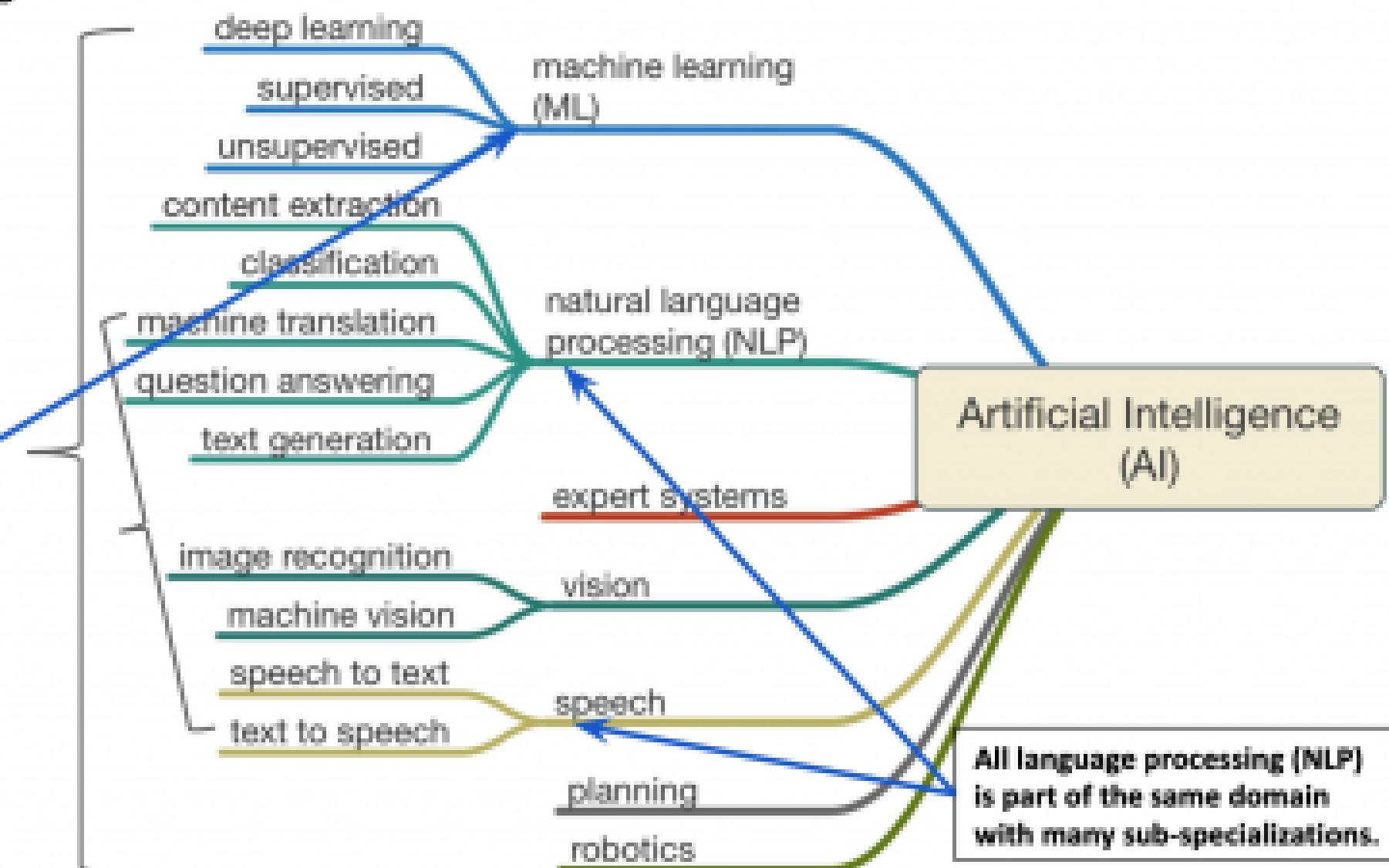


Machine Learning



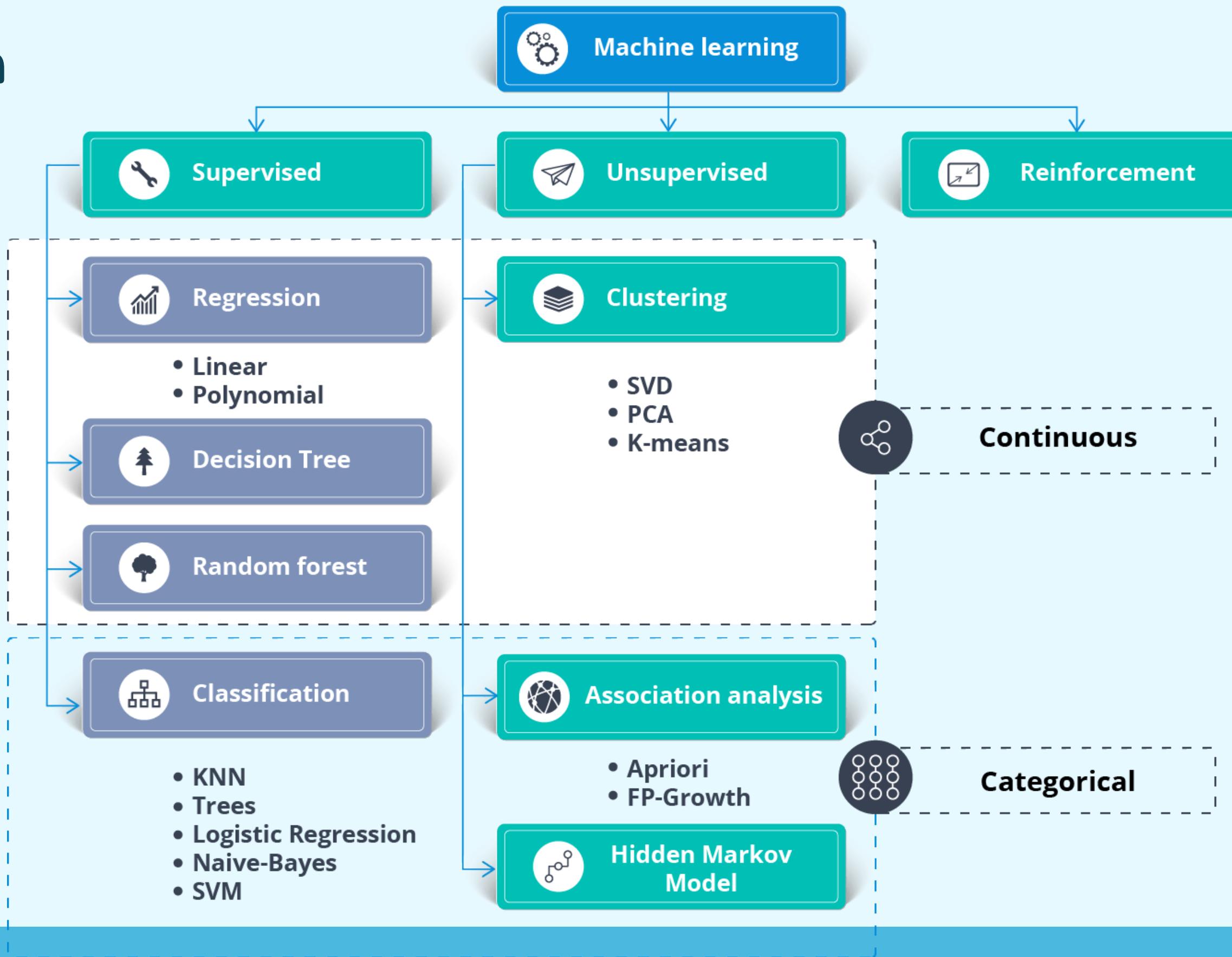
A common view of the main approaches and domains of AI /ML

"Machine Learning" and **"Deep Learning"** techniques are now being applied in the other domains.



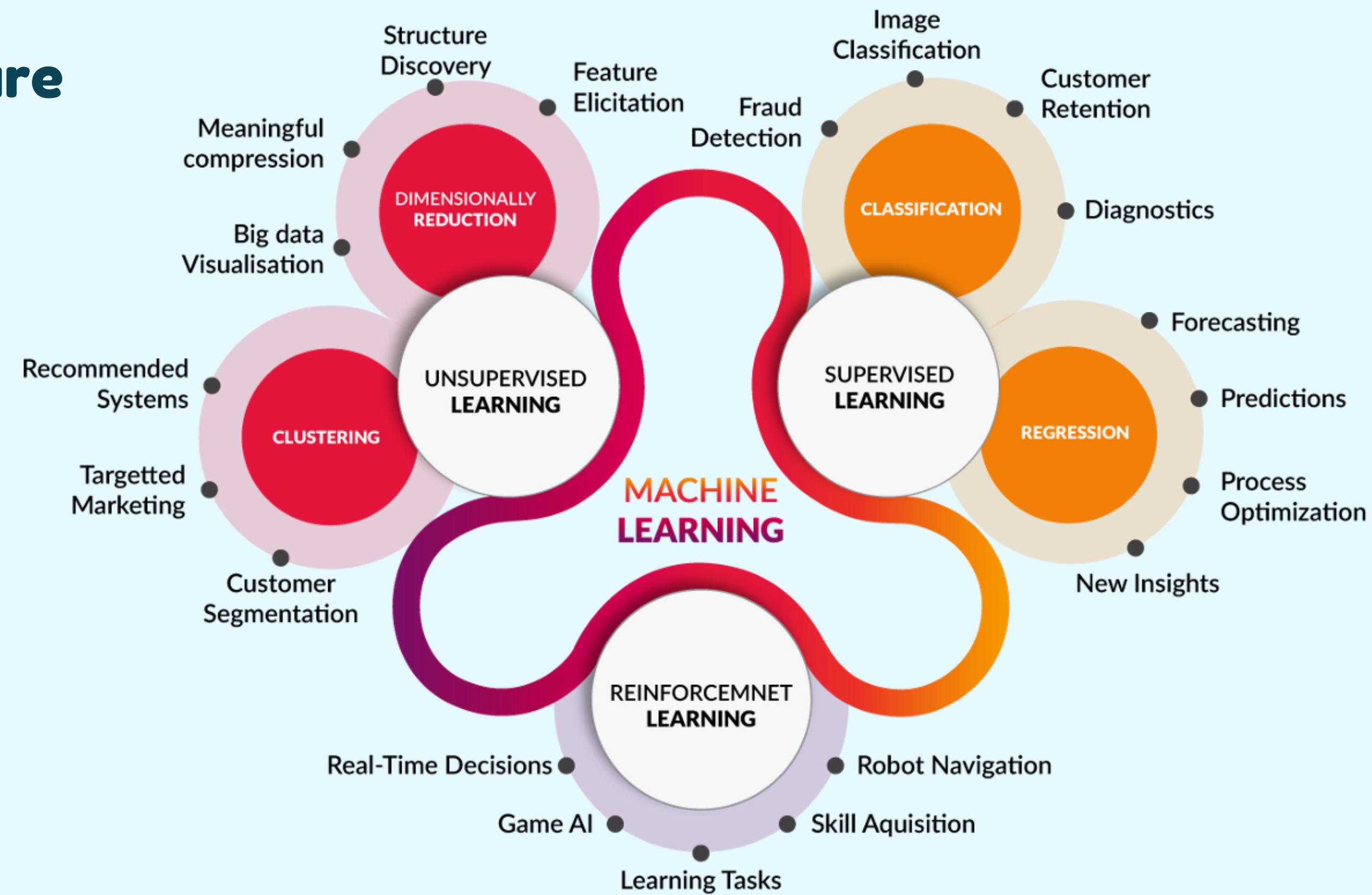


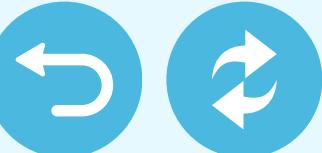
Machine Algorithm





ML Structure





Deploy your End to End projects to show your clients / users

kaggle

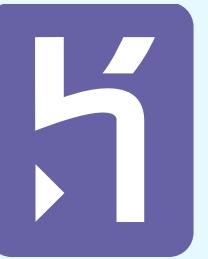
aws



Google Cloud



Microsoft
Azure



heroku

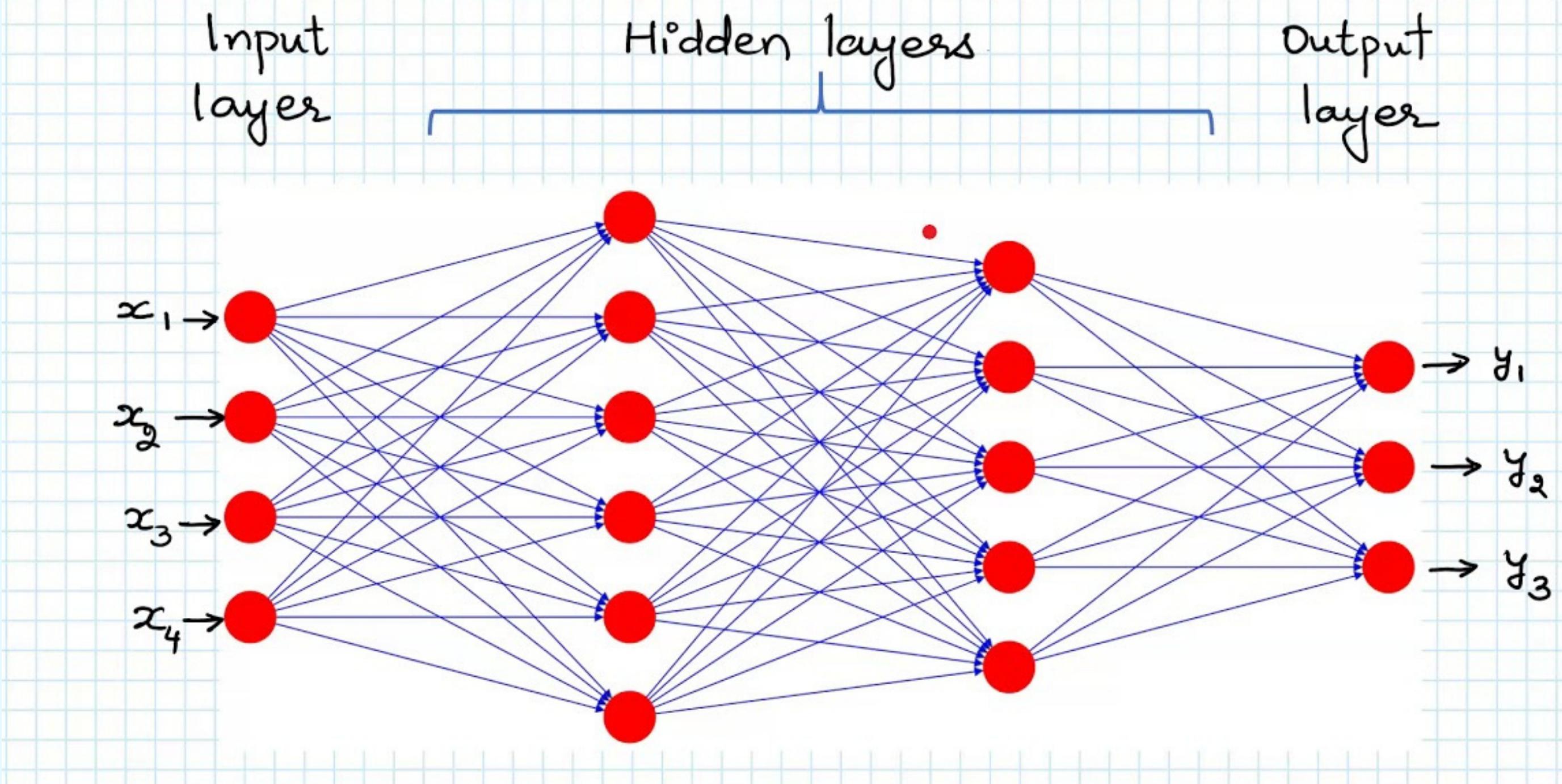


Deep learning

- ANN
- CNN
- RNN
- Transfer Learning
- Auto Encoder

ANN(Artificial Neural network)

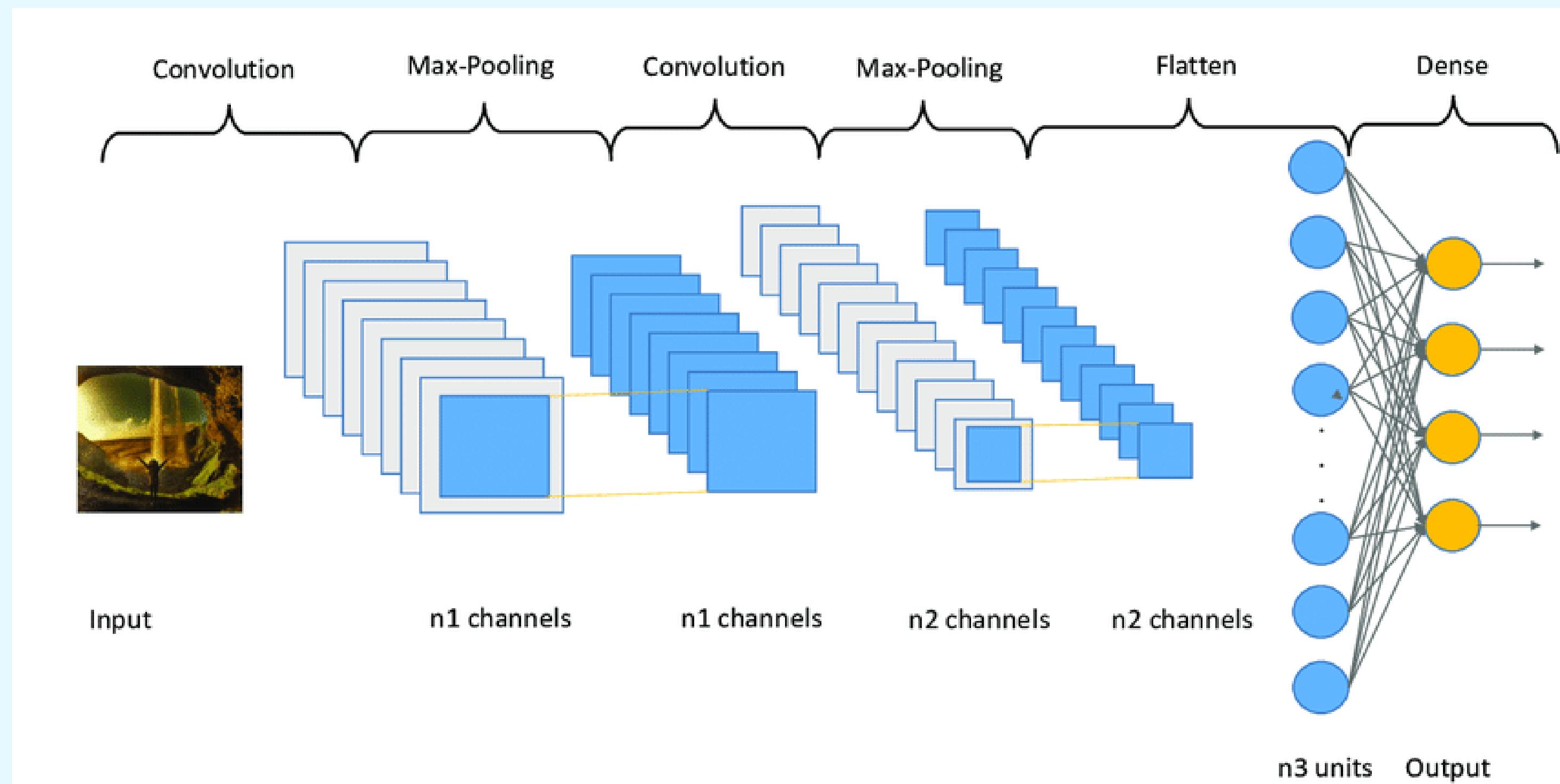
Artificial Neural Network (ANN) – Feedforward ANN





Deep learning

CNN (Convolutions Neural network)

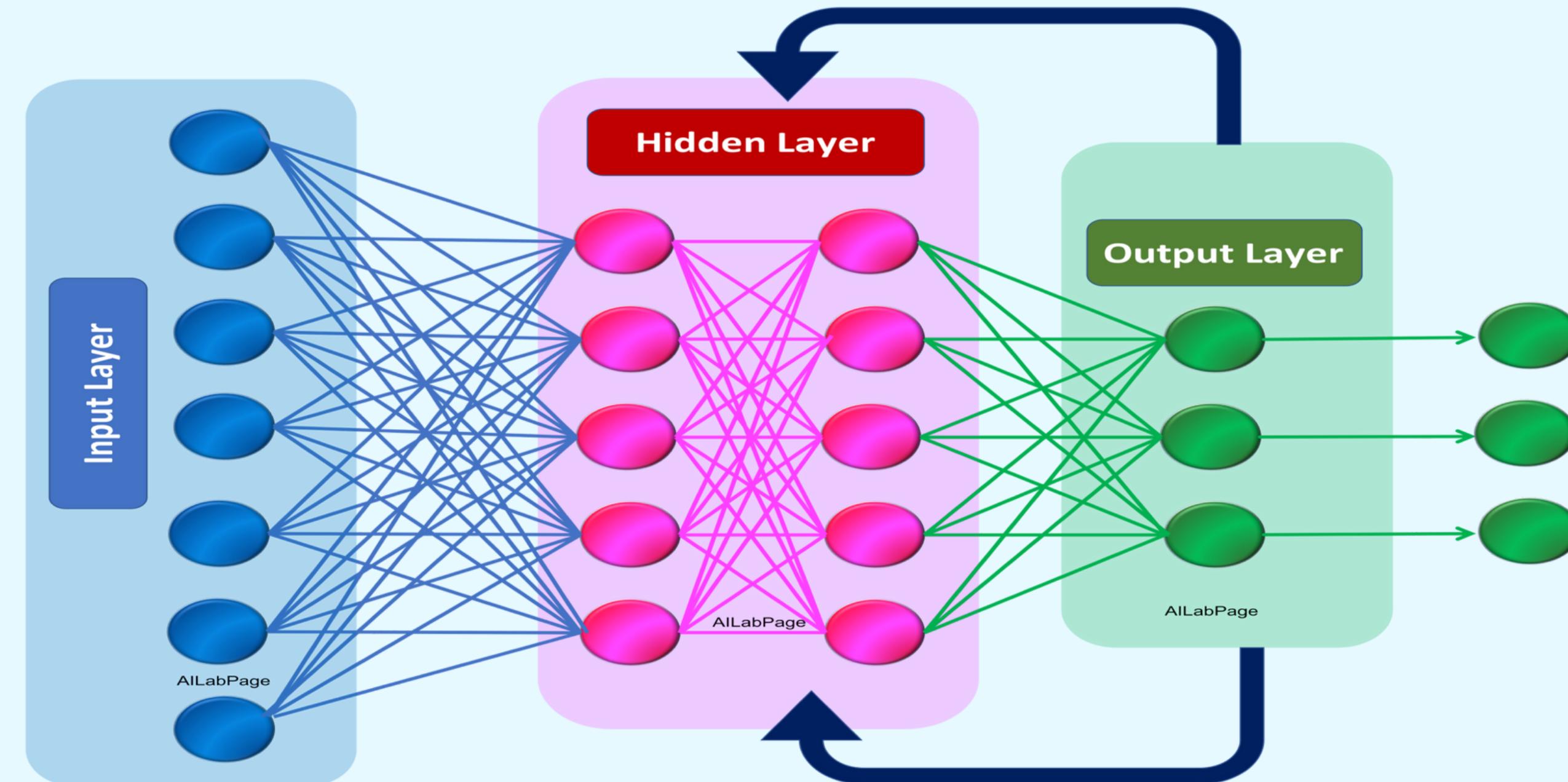




RNN (Recurrent Neural network)

Recurrent Neural Networks

Deep learning

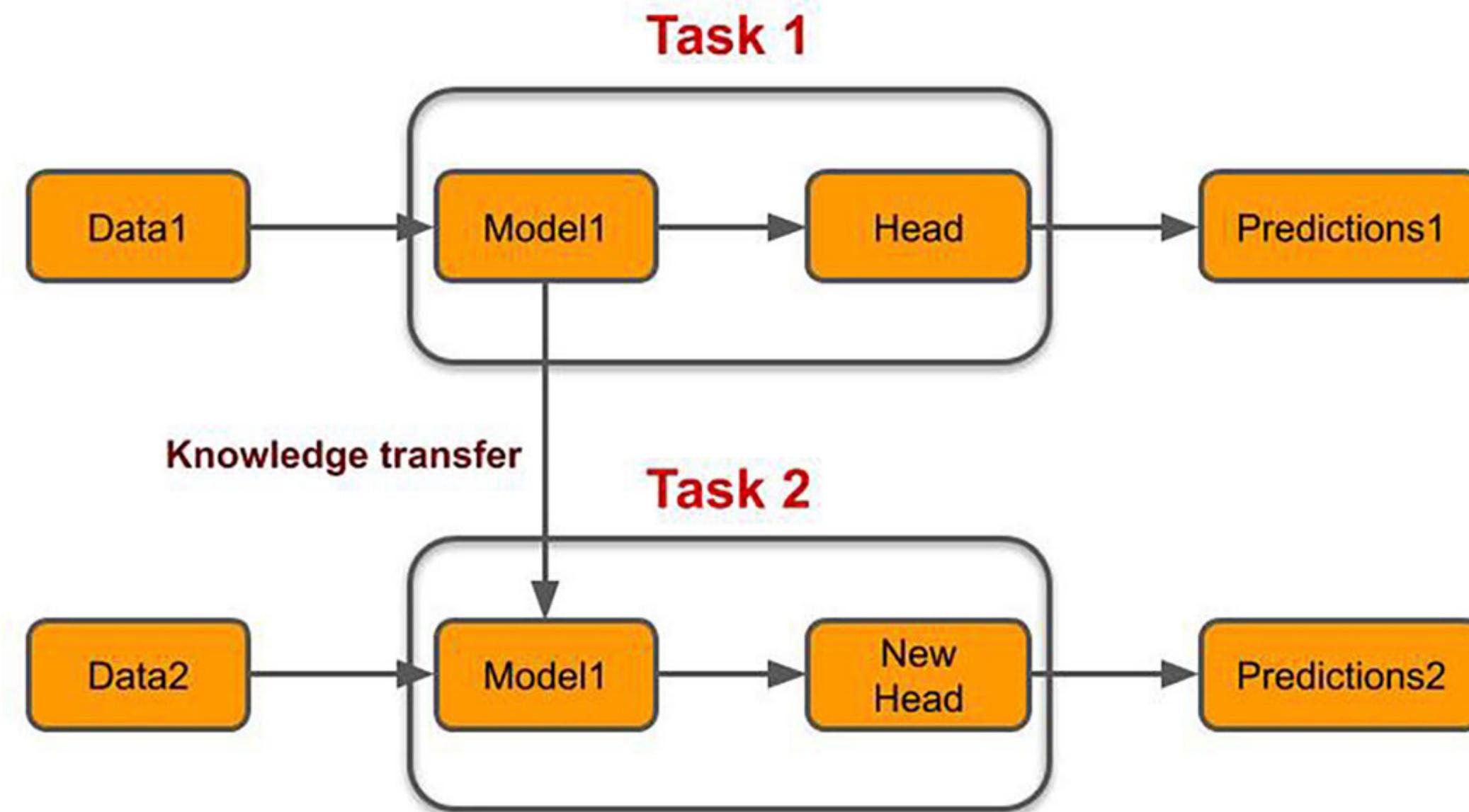




Deep learning

Transfer learning

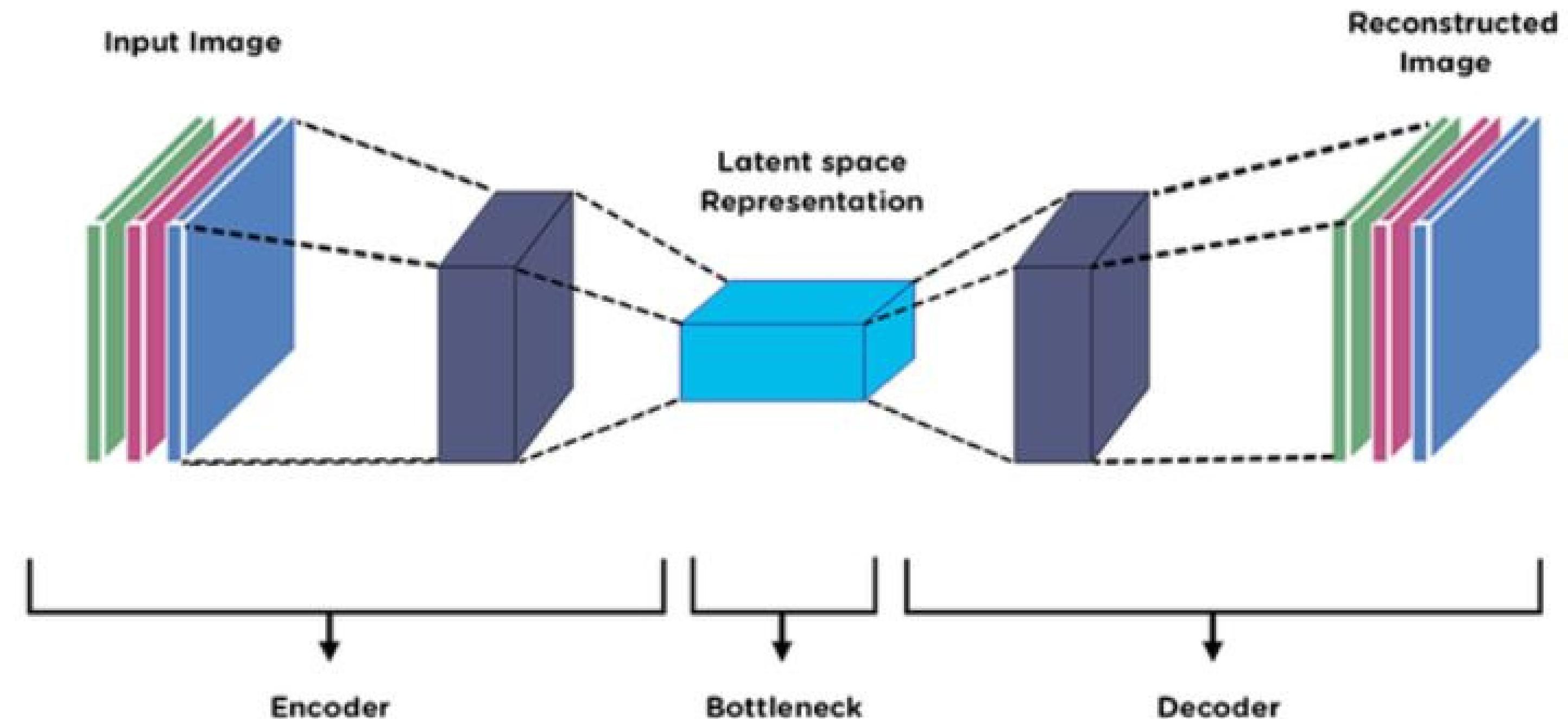
Transfer Learning





Auto Encoder

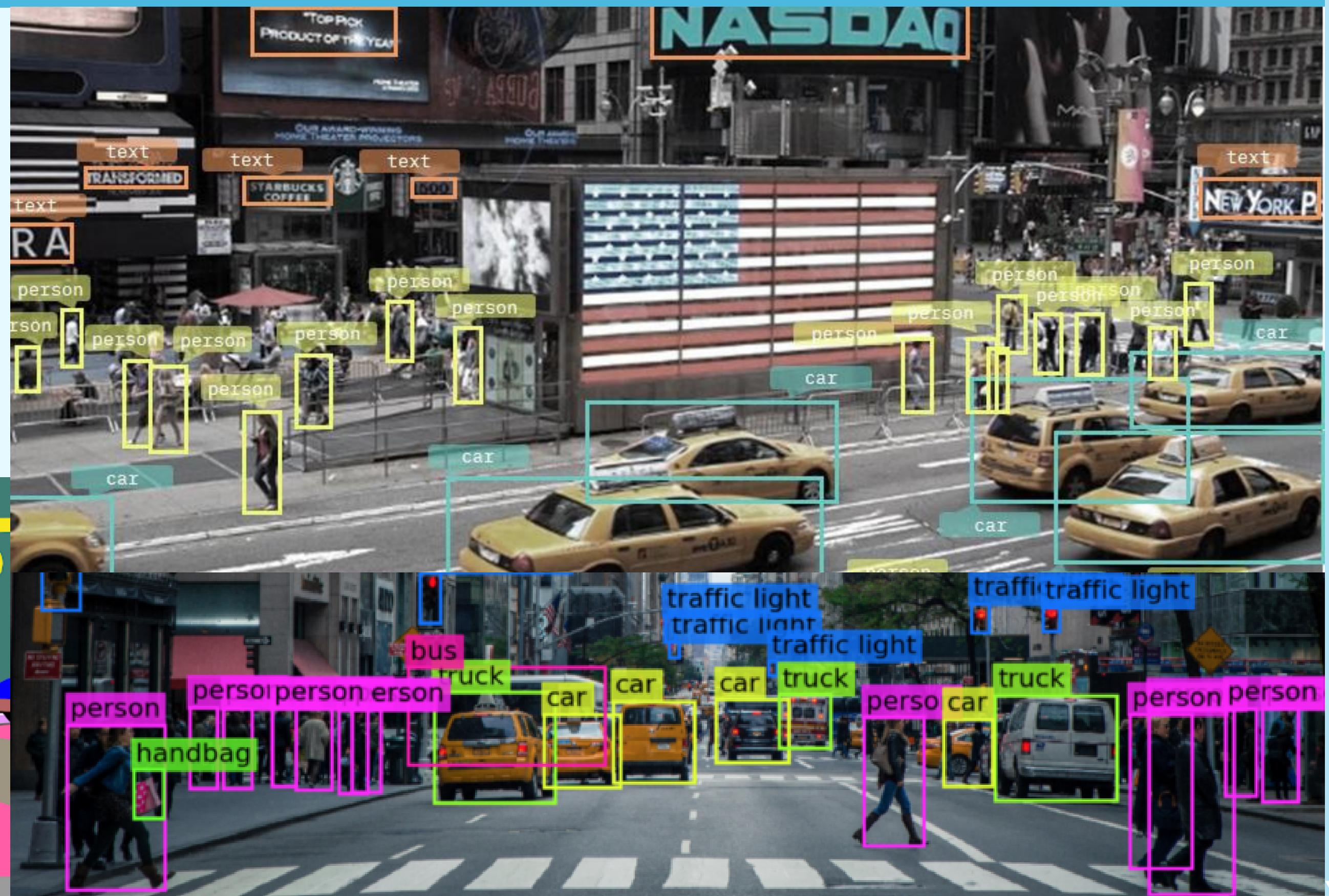
Deep learning





Computer Vision

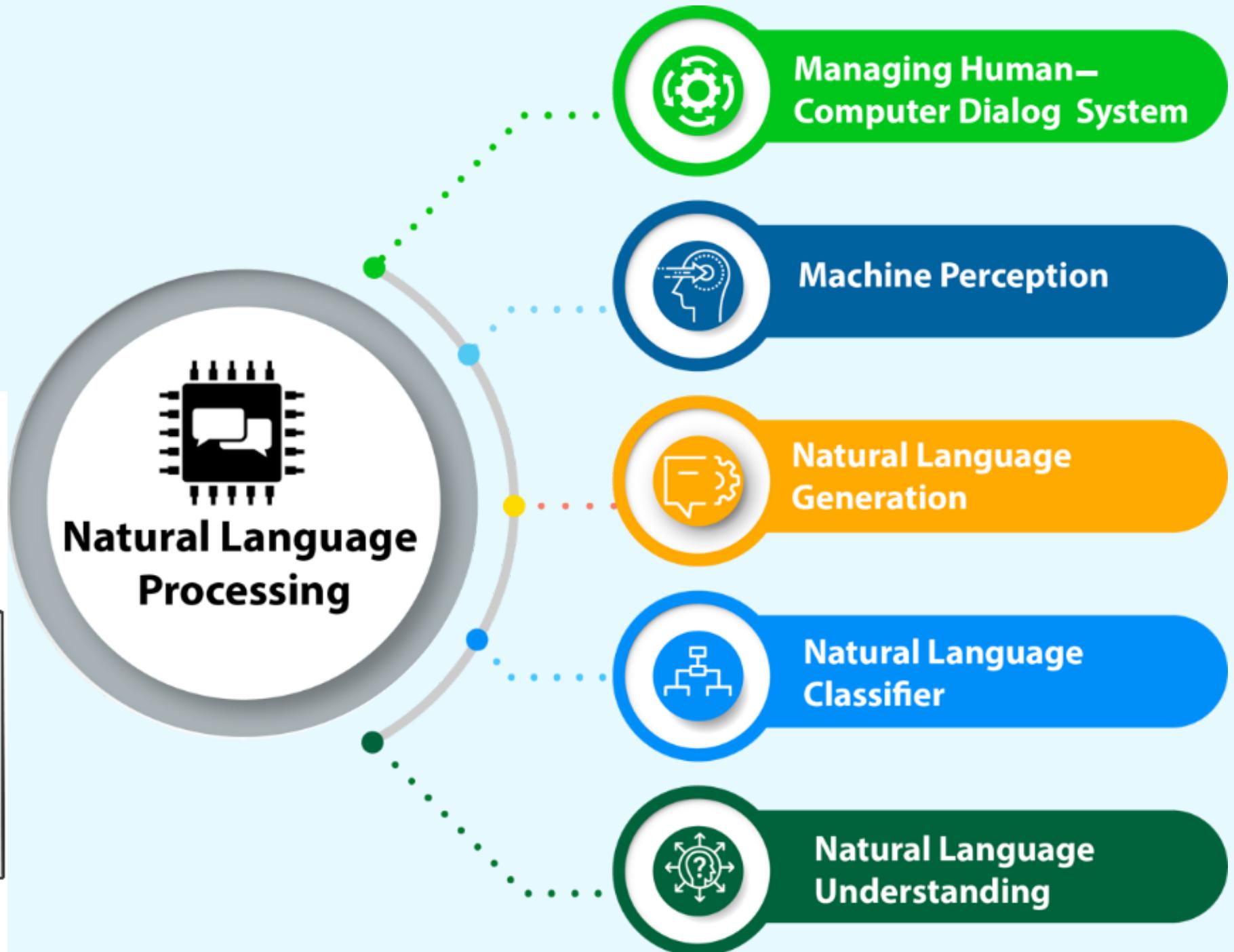
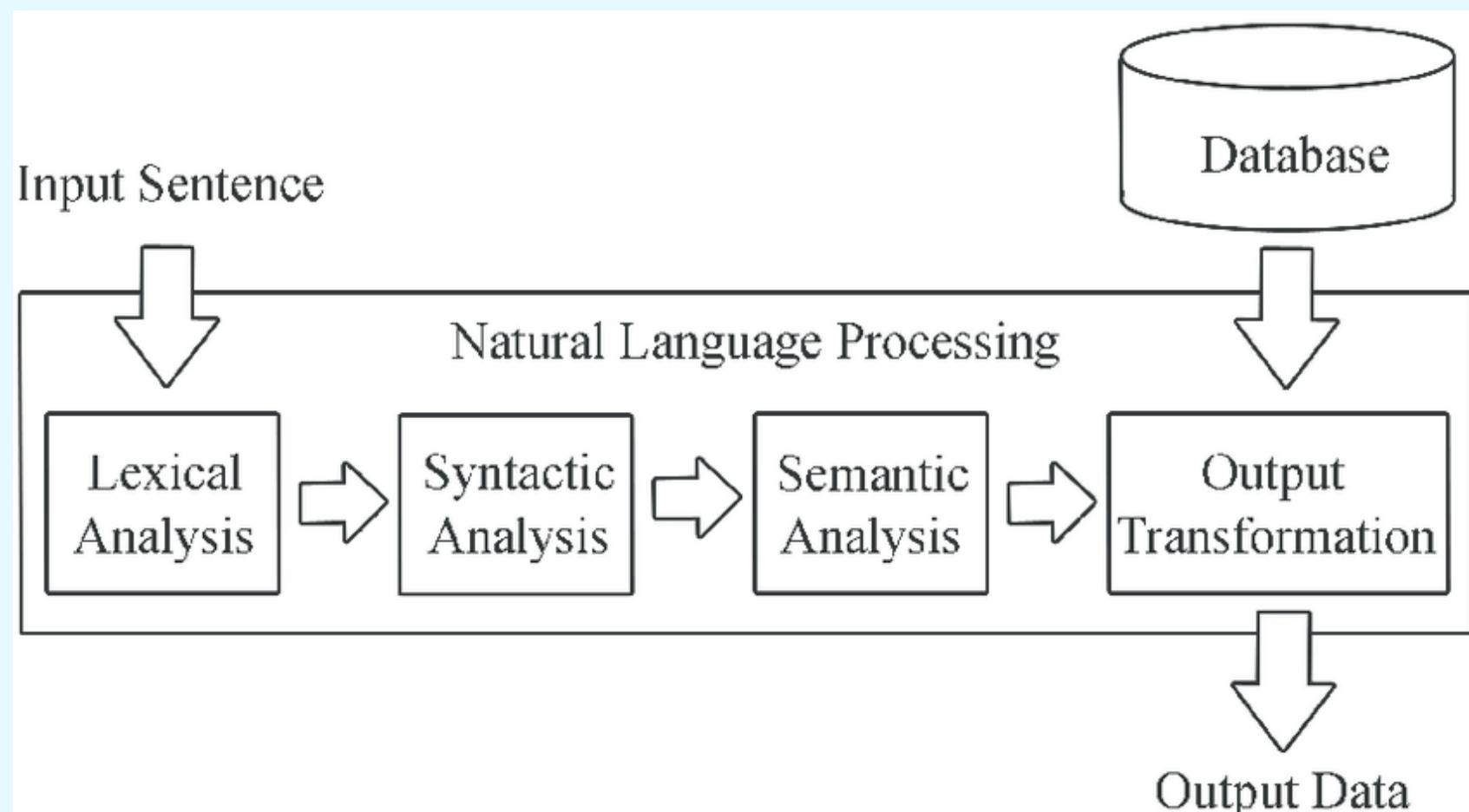
- image Classification
- Object Detection
- Semantic Segment
- Object Tracking
- Instance Segment





Natural Language Processing

- Bag of Words
- Stemming
- TF-IDF
- Lemmatization
- Tokenization
- Topic modeling
- Stop Word Removal
- Word Embedding





Thanks for watching



Thank
you!