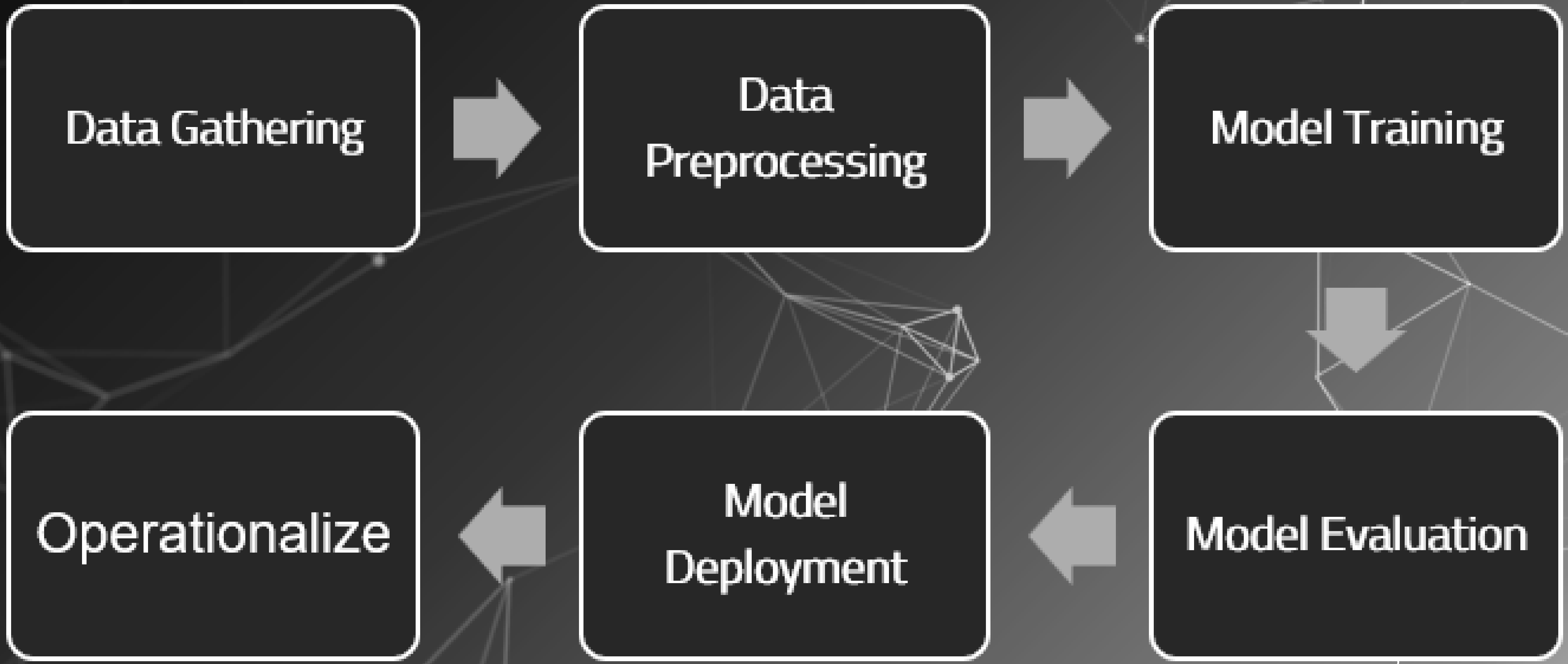
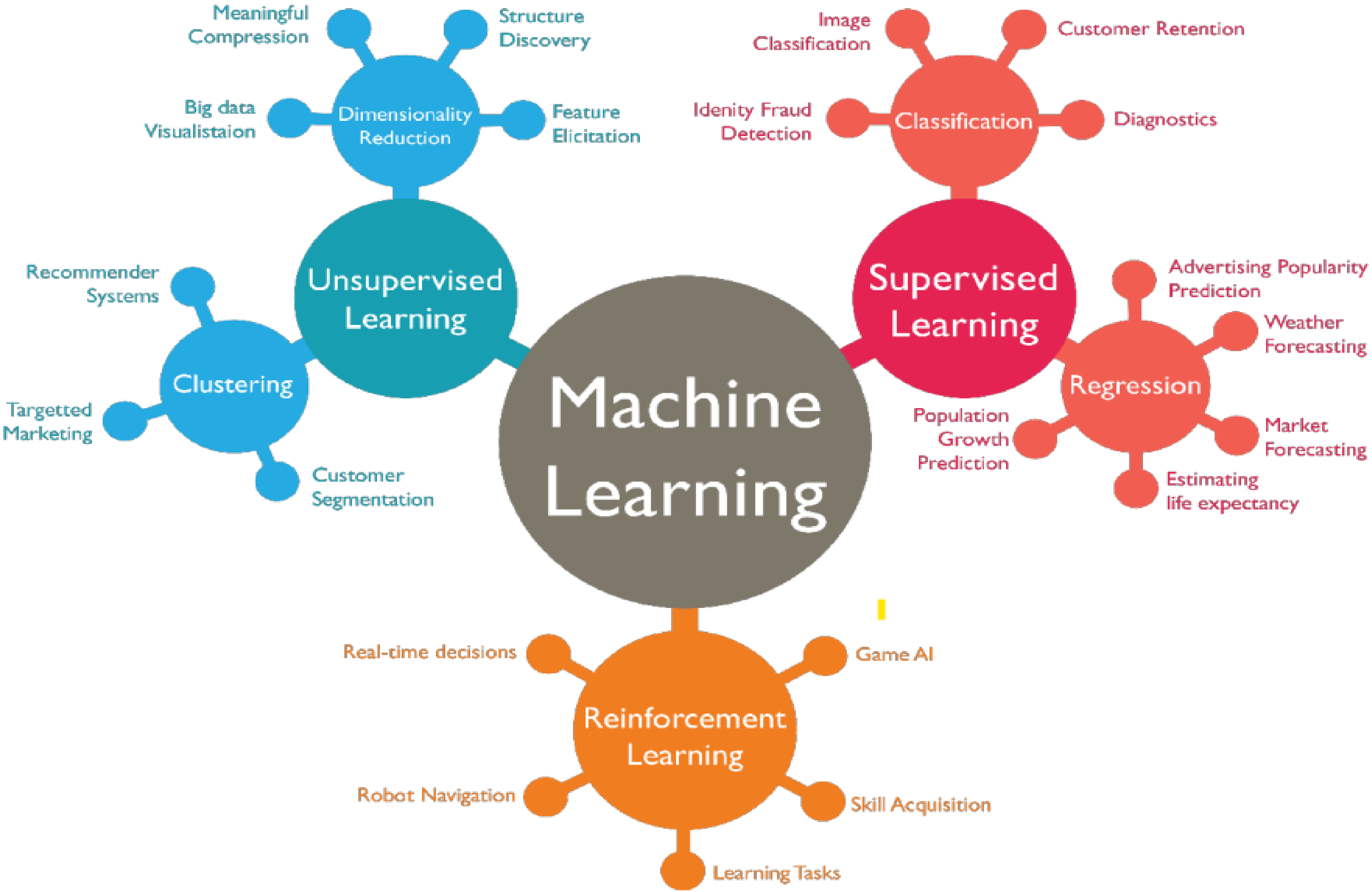
INTRO TO COMPUTER VISION & NLP

ENG- MOHAMED KHALED IDRIS

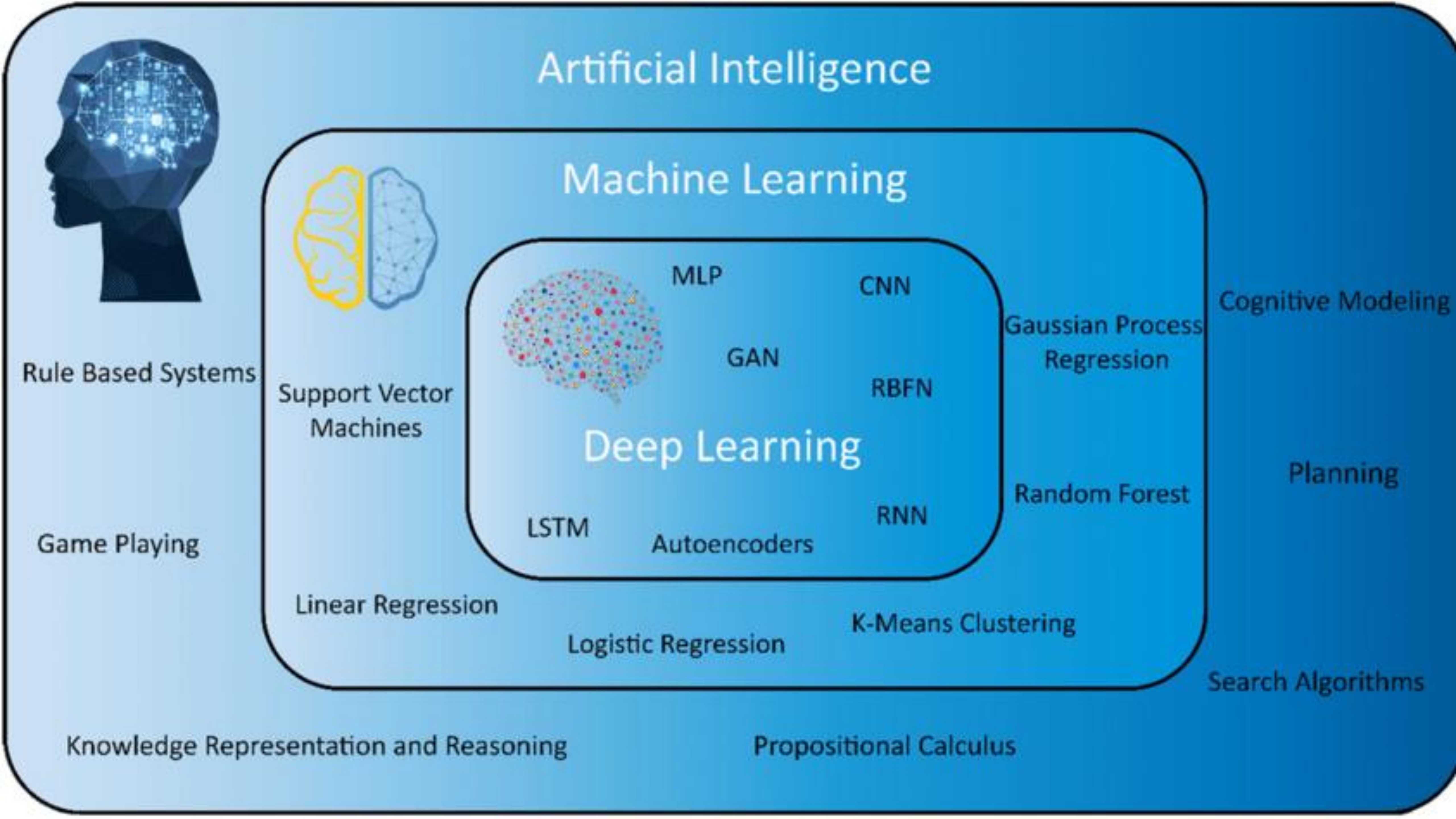
ENG- MAYAR SWILAM

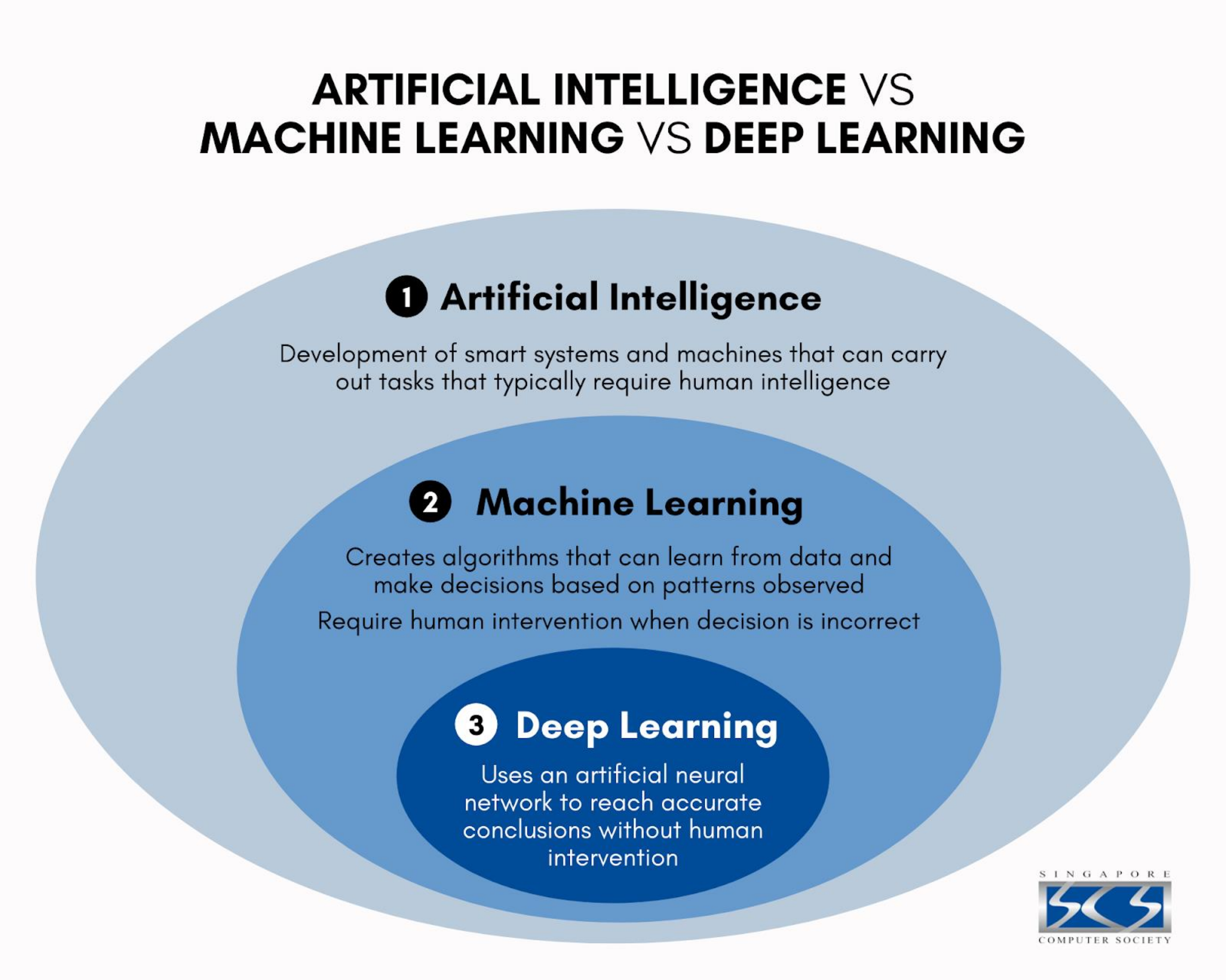
Machine Learning Process





Machine Learning
Unsupervised Learning
Dimensionality Reduction
Meaningful Compression
Structure Discovery
Big data Visualistaion
Feature Elicitation
Clustering
Recommender Systems
Targetted Marketing
Customer Segmentation
Supervised Learning
Classification
Image Classification
Customer Retention
Idenity Fraud Detection
Diagnostics
Regression
Advertising Popularity Prediction
Weather Forecasting
Population Growth Prediction
Market Forecasting
Estimating life expectancy
Reinforcement Learning
Real-time decisions
Game AI
Robot Navigation
Skill Acquisition
Learning Tasks


Artificial Intelligence
Machine Learning
MLP
CNN
GAN
RBFN
Deep Learning
LSTM
Autoencoders
RNN
Cognitive Modeling
Gaussian Process Regression
Rule Based Systems
Support Vector Machines
Planning
Random Forest
Game Playing
Linear Regression
K-Means Clustering
Logistic Regression
Search Algorithms
Knowledge Representation and Reasoning
Propositional Calculus


ARTIFICIAL INTELLIGENCE VS
MACHINE LEARNING VS DEEP LEARNING
1 Artificial Intelligence
Development of smart systems and machines that can carry out tasks that typically require human intelligence
2 Machine Learning
Creates algorithms that can learn from data and make decisions based on patterns observed
Require human intervention when decision is incorrect
3 Deep Learning
Uses an artificial neural network to reach accurate conclusions without human intervention
SINGAPORE
SCS
COMPUTER SOCIETY

DATA SCIENCE AND ITS HISTORY

- The term "Data Science" originated in the early 1960s to address the challenges of interpreting and understanding the increasing volumes of data being generated. Initially focused on supporting various fields with statistical methods, Data Science has evolved over the years to incorporate computer science, artificial intelligence, machine learning, and the Internet of Things. Statistics remains a fundamental aspect of Data Science.

DATA SCIENCE AND ITS HISTORY

1.Early Years (1960s-1980s):

1. The foundations of data science can be traced back to statistics and computer science. Statisticians have long been involved in analyzing data to draw conclusions and make predictions.
2. During the 1960s and 1970s, the advent of computers and the growth of databases led to the development of techniques for managing and processing large datasets.

2.The Rise of Big Data (1990s-2000s):

1. The term "data science" began to gain prominence in the late 1990s. However, it wasn't until the early 2000s that the concept started to crystallize.
2. With the explosion of the internet, e-commerce, and social media, vast amounts of data became available, giving rise to the era of big data. Companies and researchers started grappling with the challenges of handling and extracting value from massive datasets.

3.Technological Advancements:

1. Parallel processing, distributed computing, and advancements in storage technologies played a crucial role in enabling the processing of large-scale datasets. Technologies like Hadoop emerged to handle distributed data processing

DATA SCIENCE AND ITS HISTORY

5.Emergence of Data Science as a Discipline (2010s):

1. As the need for expertise in handling big data grew, organizations began to recognize the importance of professionals who could not only manage data but also analyze it to extract actionable insights.
2. Data science became an interdisciplinary field, borrowing concepts from statistics, mathematics, computer science, and domain expertise. The ability to derive insights from complex datasets became a valuable skill.

6. Machine Learning and Artificial Intelligence (AI):

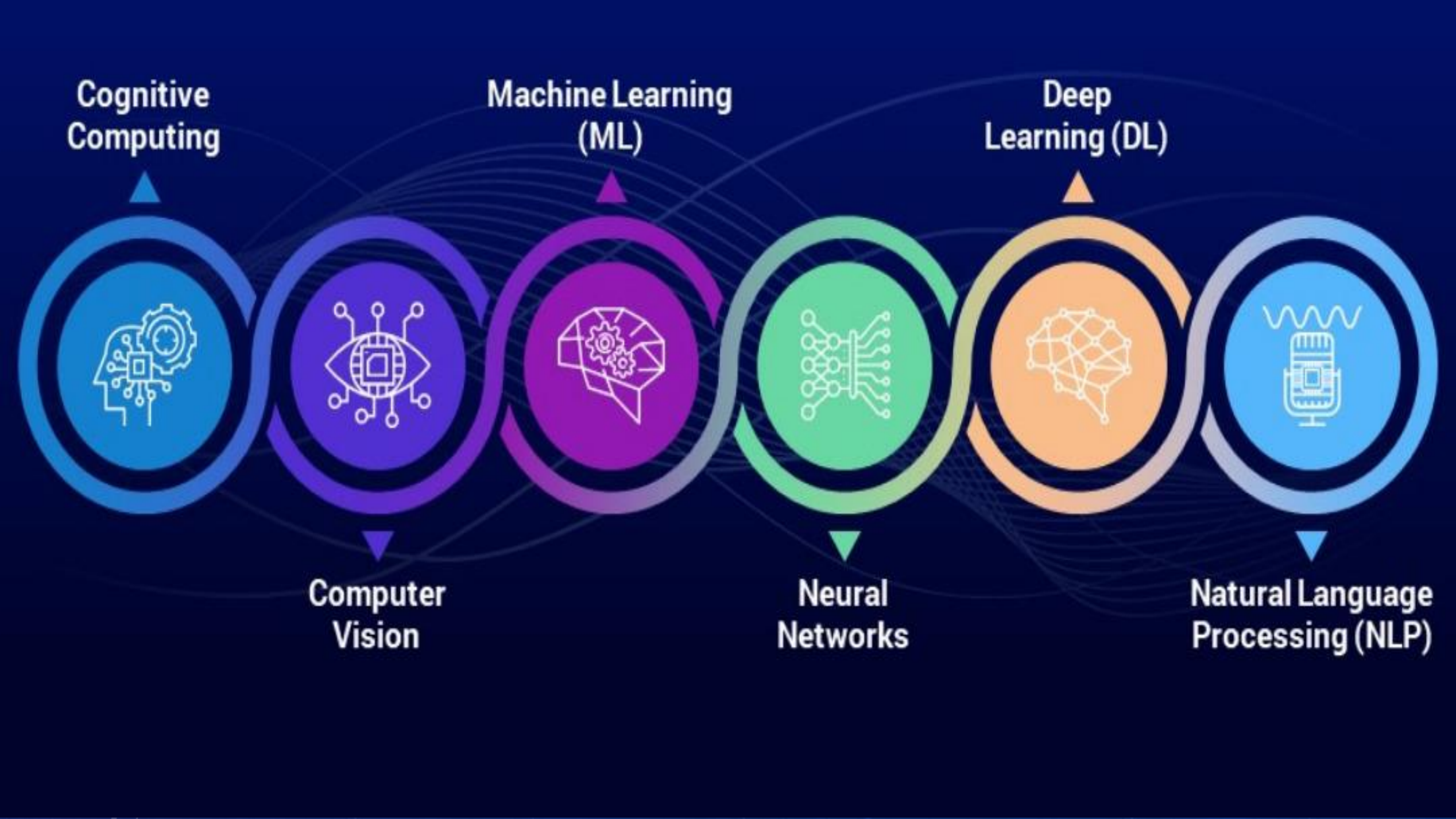
1. The integration of machine learning and AI techniques into data science further expanded its capabilities. Predictive modelling, classification, clustering, and other ML techniques became essential tools for data scientists.

7. Data Science in Industry:

1. Organizations across various sectors, including finance, healthcare, marketing, and more, started employing data scientists to gain a competitive advantage, improve decision-making, and optimize processes.

8. Ongoing Evolution (2020s and Beyond):

1. Data science continues to evolve with advancements in technology, including the growth of cloud computing, the adoption of advanced analytics tools, and the integration of artificial intelligence into various applications.


Cognitive Computing
Machine Learning (ML)
Deep Learning (DL)
Computer Vision
Neural Networks
Natural Language Processing (NLP)

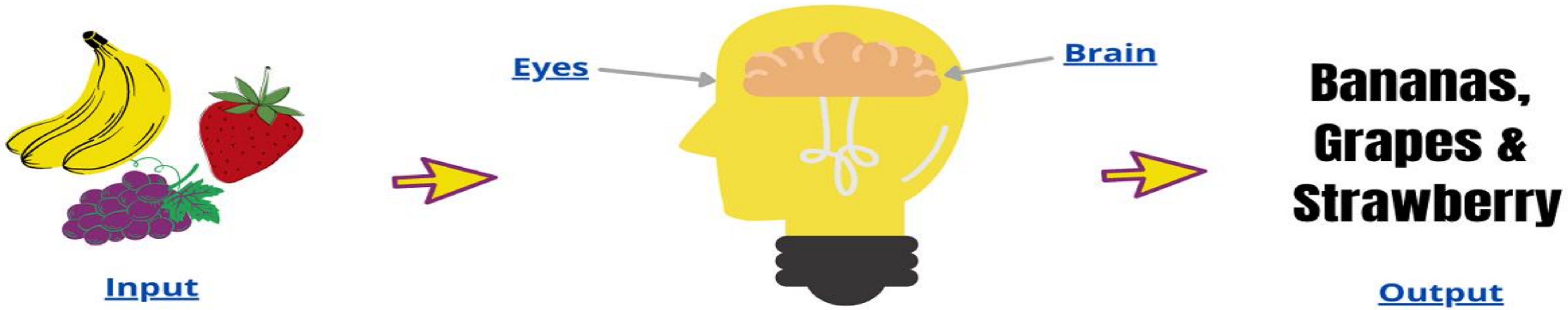
Computer Vision & NLP

Computer Vision

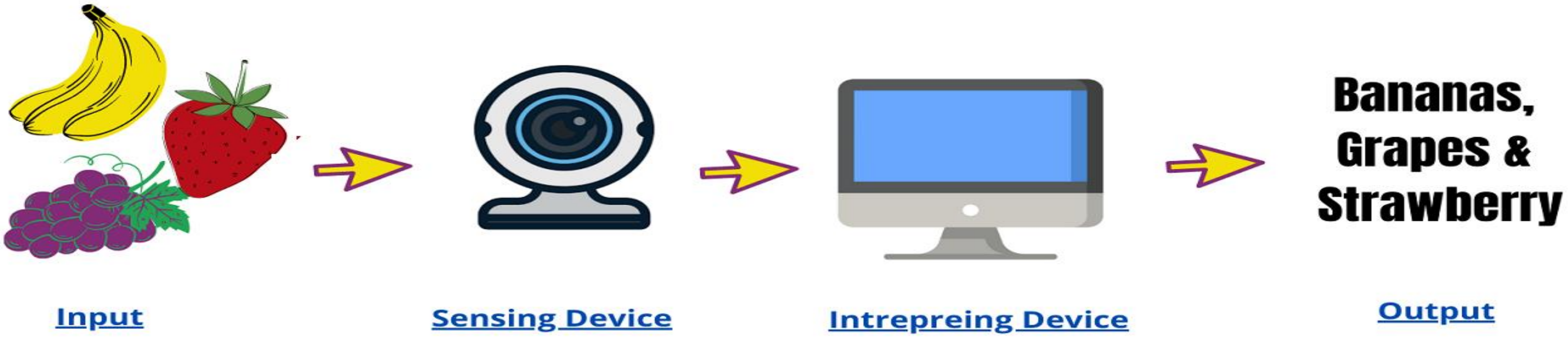
Computer vision is a multidisciplinary field that enables computers to interpret and understand visual information from the world, similar to how humans perceive and interpret visual data. It involves the development of algorithms, models, and systems that allow machines to analyze and make decisions based on images or videos.

Computer vision applications are diverse and can be found in various industries, including healthcare, automotive, agriculture, security, and entertainment.

Human Vision System



Computer Vision System



Core Concepts in Computer Vision

IMAGE CLASSIFICATION

It involves categorizing an image into a set of predefined classes such as dog, cat, car, etc. Traditional approaches relied on extracting handcrafted features like SIFT or HOG and training classifiers like SVM.

OBJECT DETECTION

The goal object detection is to not only classify objects but also localize them within an image by drawing bounding boxes around them. Region-based CNNs like R-CNN first generate object proposals then classify each region.

SEMANTIC SEGMENTATION

This associates each pixel in an image with a class label like person, car, road, etc. Fully convolutional networks are commonly used which classify each pixel independently in an end-to-end manner.

INSTANCE SEGMENTATION

his identifies and delineates each distinct object instance in an image. Mask R-CNN extends object detection networks to also generate a segmentation mask specific to each detected object instance

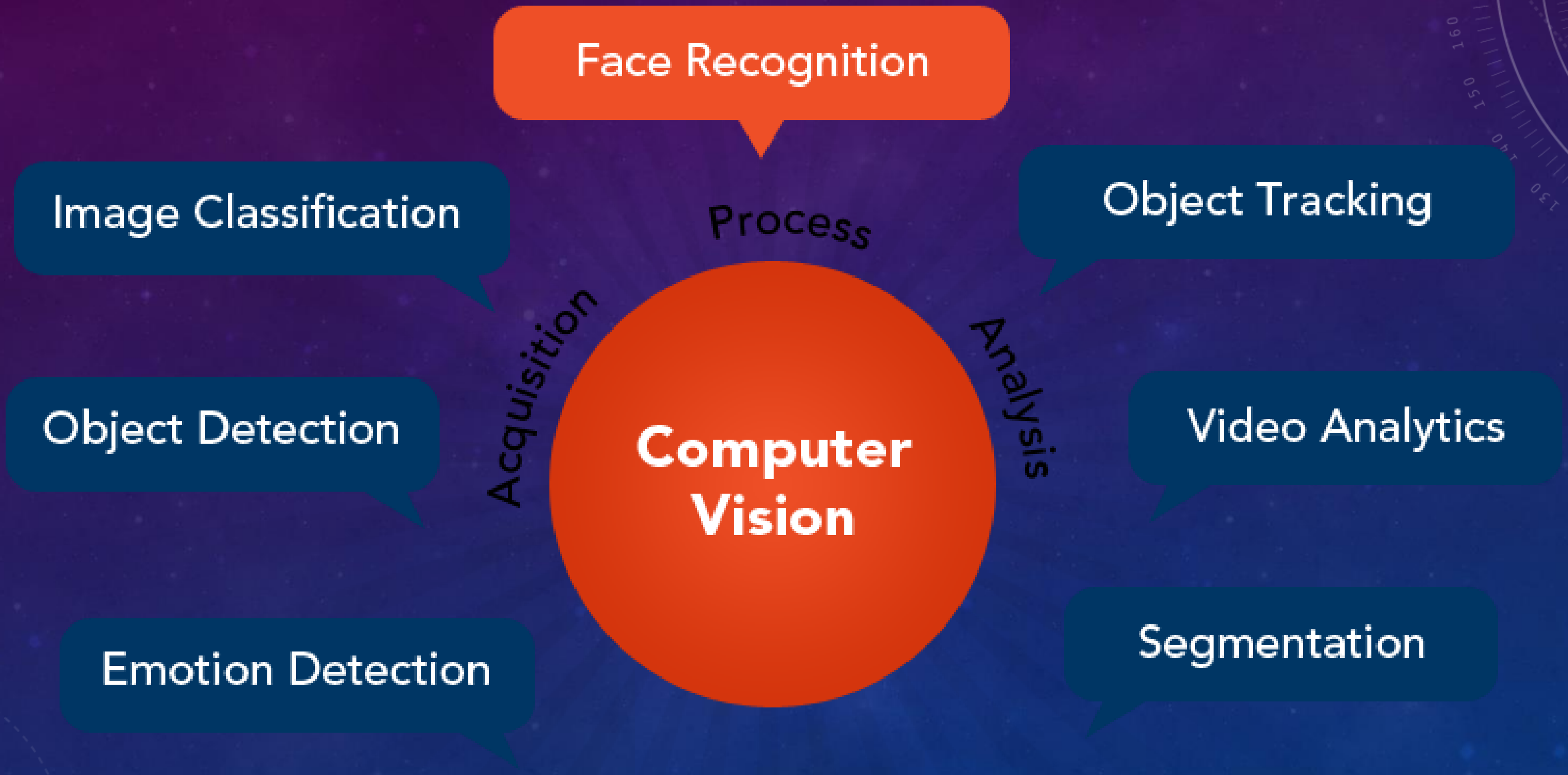
FEATURE EXTRACTION AND MATCHING

Distinctive feature points in images like corners and edges can be extracted using algorithms such as SIFT and SURF. Robust descriptors are computed at each feature point.

3D COMPUTER VISION

Reconstructing 3D models and structures from 2D images is an important problem. Using feature matching across stereo images from multiple viewpoints helps estimating depth and 3D structure via triangulation.

Computer Vision




OpenCV
K
computer vision libraries

Computer Vision Libraries

1.OpenCV (Open Source Computer Vision Library):

1. OpenCV is a widely-used open-source computer vision library that provides a vast collection of tools and algorithms for image and video processing, feature extraction, object detection, machine learning, and more. It supports multiple programming languages, including Python, C++, and Java.

2.TensorFlow:

1. Developed by Google, TensorFlow is an open-source machine learning library that includes modules for deep learning and neural networks. TensorFlow provides a high-level API called TensorFlow Keras for building and training neural networks, making it suitable for computer vision tasks.

3.PyTorch:

1. PyTorch is an open-source deep learning library developed by Facebook. It has gained popularity for its dynamic computational graph and ease of use. PyTorch is commonly used for research in computer vision and natural language processing. It also provides the torchvision module for computer vision-specific functionalities.

Computer Vision Libraries

1.Keras:

1. Keras is a high-level neural networks API written in Python and capable of running on top of TensorFlow, Theano, or Microsoft Cognitive Toolkit (CNTK). While it's often used for general deep learning tasks, it also includes modules for image classification, object detection, and other computer vision applications.

2.Scikit-Image:

1. Scikit-Image is a collection of algorithms for image processing in Python. It builds on the capabilities of NumPy and provides tools for tasks such as filtering, morphology, segmentation, and feature extraction.

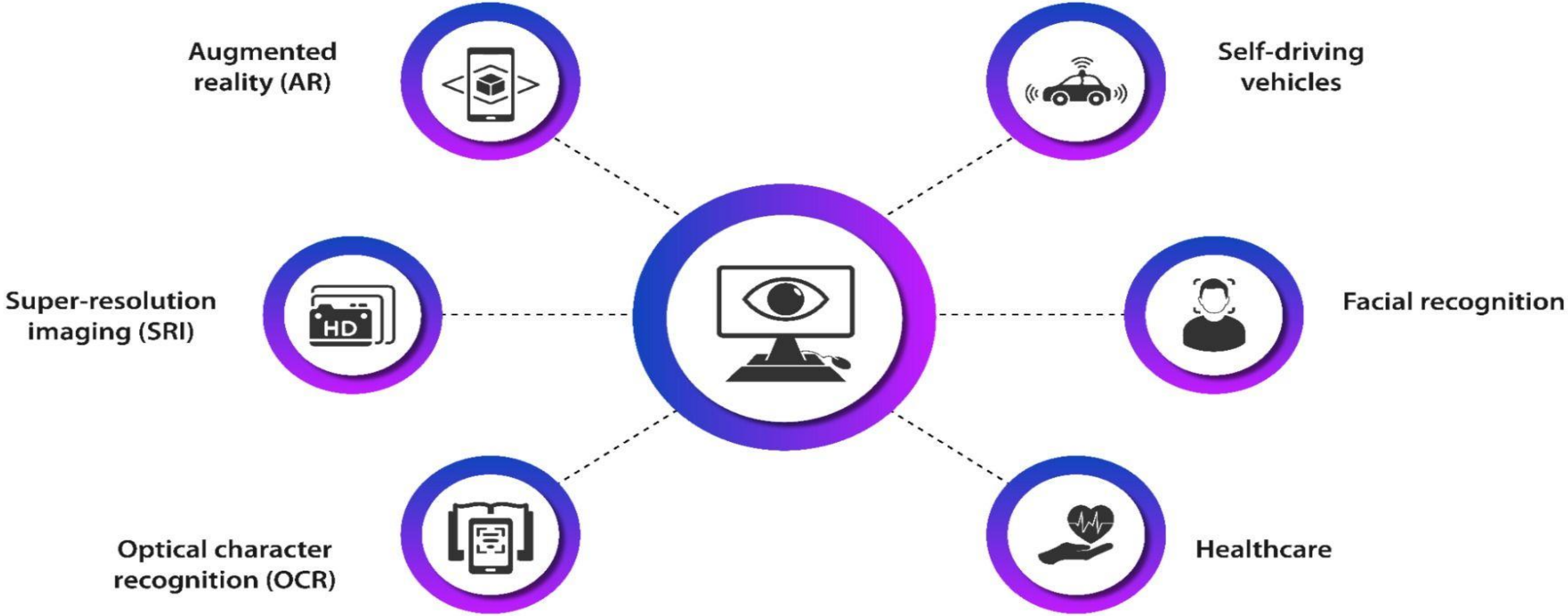
3.Dlib:

1. Dlib is a C++ toolkit for machine learning and computer vision. It includes a variety of algorithms for image processing, facial recognition, object detection, and more. Dlib is known for its efficiency and performance.

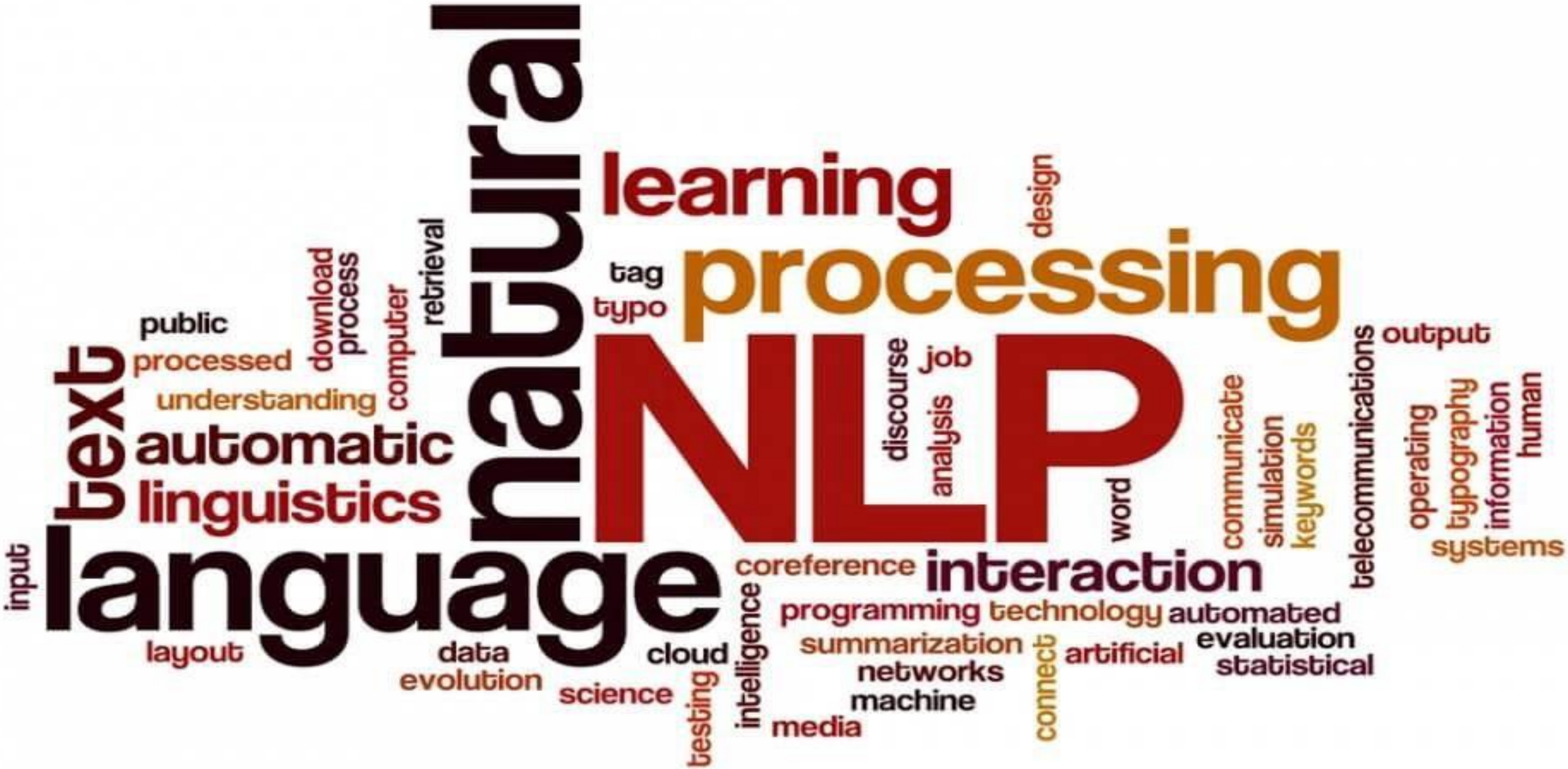
Computer Vision Libraries

1. **MXNet:**
    1. Apache MXNet is an open-source deep learning framework designed for efficiency and scalability. It supports both symbolic and imperative programming and includes GluonCV, a toolkit for computer vision tasks built on top of MXNet.
2. **Caffe:**
    1. Caffe is a deep learning framework developed by the Berkeley Vision and Learning Center. It is popular for its speed and efficiency in convolutional neural networks. While it's not as flexible as some other frameworks, it excels in image classification tasks.
3. **Fastai:**
    1. Fastai is a high-level deep learning library built on top of PyTorch. It aims to make deep learning more accessible through high-level abstractions and simplifications. Fastai includes modules for computer vision, natural language processing, and tabular data.
4. **MATLAB Computer Vision Toolbox:**
    1. MATLAB provides a Computer Vision Toolbox that offers a comprehensive set of functions for image and video processing, computer vision, and machine learning. It is widely used in academic and research environments.

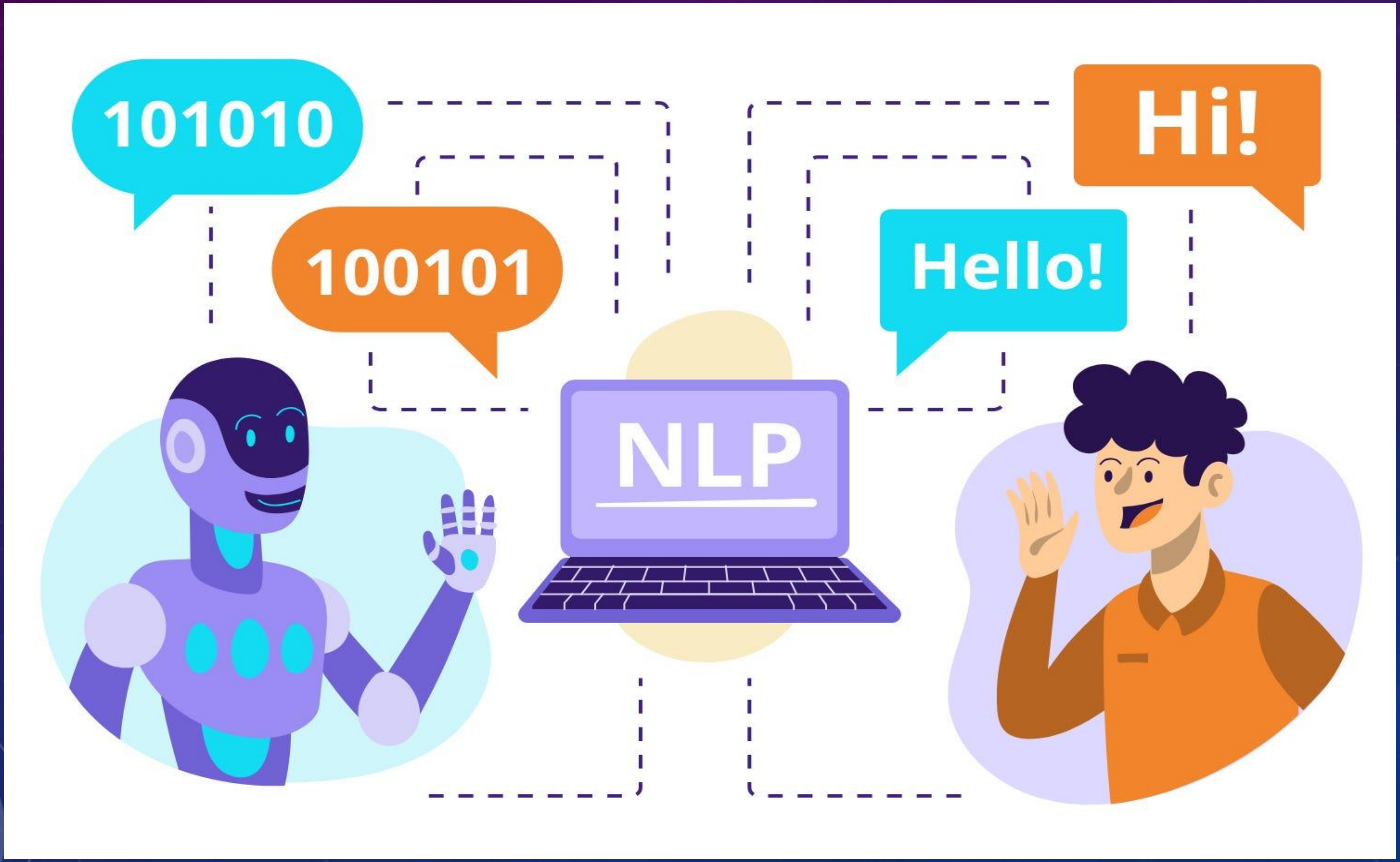
Computer Vision Applications



Turing


natural
learning
processing
NLP
language
text
design
retrieval
download
process
computer
tag
typo
public
processed
understanding
automatic
linguistics
discourse
job
analysis
word
communicate
simulation
keywords
telecommunications
output
operating
typography
information
human
systems
input
coreference
interaction
programming
technology
automated
evaluation
statistical
layout
data
evolution
cloud
intelligence
summarization
connect
artificial
networks
machine
science
testing
media

Natural language processing (NLP)



Natural language processing (NLP)

Natural Language Processing (NLP) is a subfield of artificial intelligence (AI) that focuses on the interaction between computers and human language. The goal of NLP is to enable computers to understand, interpret, and generate human language in a way that is both meaningful and contextually relevant. It involves the development of algorithms and models that allow computers to perform tasks such as text and speech recognition, language translation, sentiment analysis, and more.

Natural language processing Algorithms

1. Tokenization:
   1. **Algorithm:** Basic tokenization involves breaking down a text into individual words or tokens.
   2. **Example:** Splitting the sentence "The quick brown fox jumps over the lazy dog" into individual words: ["The", "quick", "brown", "fox", "jumps", "over", "the", "lazy", "dog"].

2. Part-of-Speech Tagging (POS):
   1. **Algorithm:** Assigning grammatical parts of speech (e.g., noun, verb, adjective) to each word in a sentence.
   2. **Example:** Tagging the sentence "The cat is sleeping" with parts of speech: [("The", "Determiner"), ("cat", "Noun"), ("is", "Verb"), ("sleeping", "Verb")].

3. Named Entity Recognition (NER):
   1. **Algorithm:** Identifying and classifying entities (e.g., names of people, organizations, locations) in a text.
   2. **Example:** Extracting entities from the sentence "Apple Inc. is headquartered in Cupertino, California": [("Apple Inc.", "Organization"), ("Cupertino, California", "Location")].

Natural language processing Algorithms

1. **Text Classification:**
    1. **Algorithm:** Assigning predefined categories or labels to a document or text.
    2. **Example:** Classifying emails as spam or non-spam based on their content.
2. **Word Embeddings:**
    1. **Algorithm:** Learning vector representations (embeddings) for words in a way that captures semantic relationships.
    2. **Example:** Word2Vec, GloVe, and FastText are popular word embedding algorithms.
3. **Sequence-to-Sequence Models:**
    1. **Algorithm:** Transforming input sequences into output sequences, often used in machine translation and summarization.
    2. **Example:** Recurrent Neural Networks (RNNs) and Transformer models like GPT and BERT.

Natural language processing Algorithms

1. **Sentiment Analysis:**
   1. **Algorithm:** Determining the sentiment or emotion expressed in a piece of text.
   2. **Example:** Classifying product reviews as positive, negative, or neutral.
2. **Dependency Parsing:**
   1. **Algorithm:** Analyzing the grammatical structure of sentences to determine the relationships between words.
   2. **Example:** Representing the syntactic structure of a sentence using a dependency tree.
3. **Machine Translation:**
   1. **Algorithm:** Automatically translating text from one language to another.
   2. **Example:** Using statistical models or neural machine translation models like Transformer architectures.

Example: English to French Translation

Step 1: Data Preparation
Let's consider a small dataset of English and French sentence pairs:
•English: "I love natural language processing."
•French: "J'adore le traitement du langage naturel."
Step 2: Tokenization
Tokenize the input (English) and output (French) sentences into individual words:
•English: ["I", "love", "natural", "language", "processing"]
•French: ["J'adore", "le", "traitement", "du", "langage", "naturel"]
Step 3: Vocabulary
Create separate vocabularies for English and French:
•English Vocabulary: ["I", "love", "natural", "language", "processing"]
•French Vocabulary: ["J'adore", "le", "traitement", "du", "langage", "naturel"]

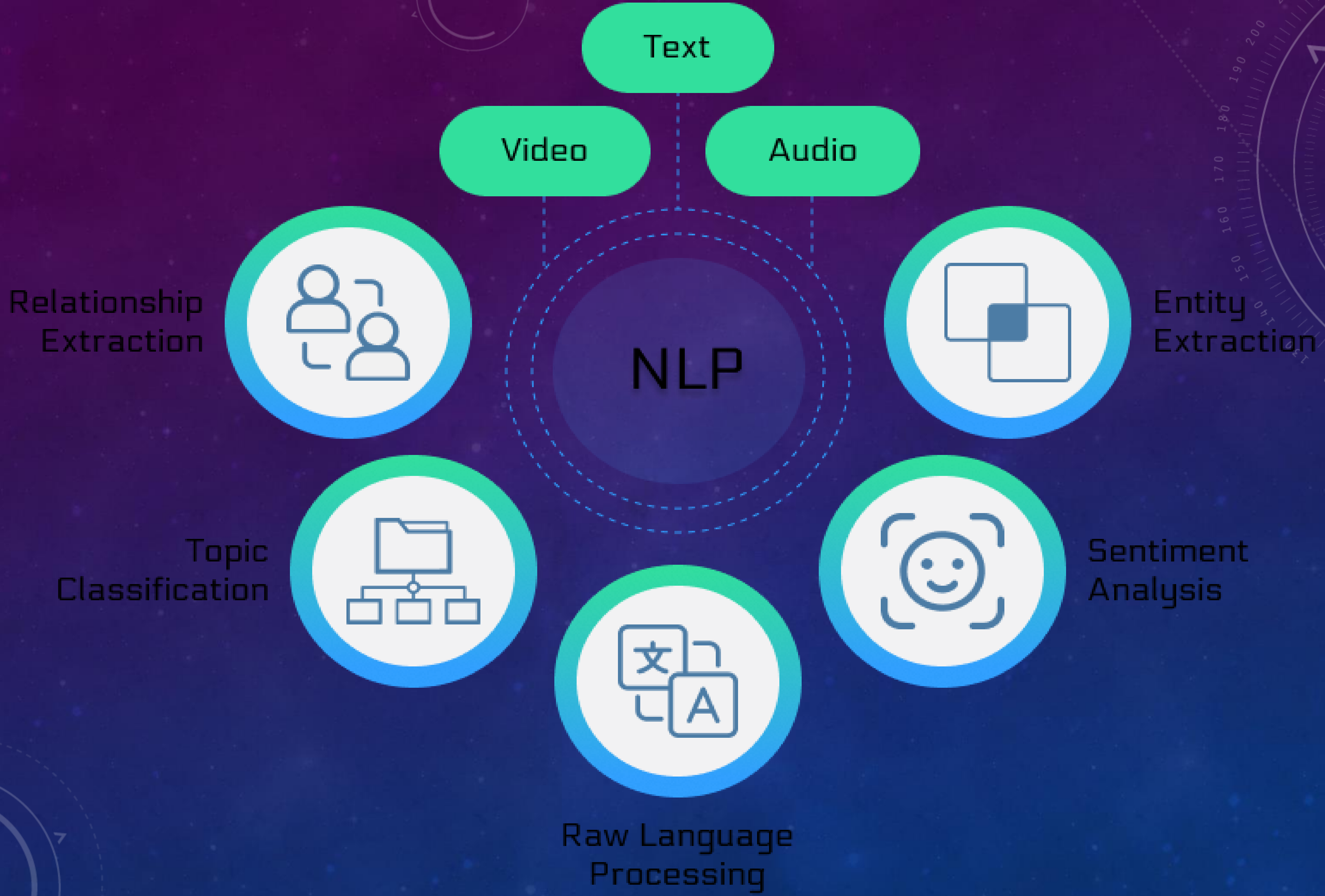
Example: English to French Translation

Step 4: Assigning Indices
Assign indices to each word in the vocabularies:
- English Indices: [0, 1, 2, 3, 4]
- French Indices: [5, 6, 7, 8, 9, 10]

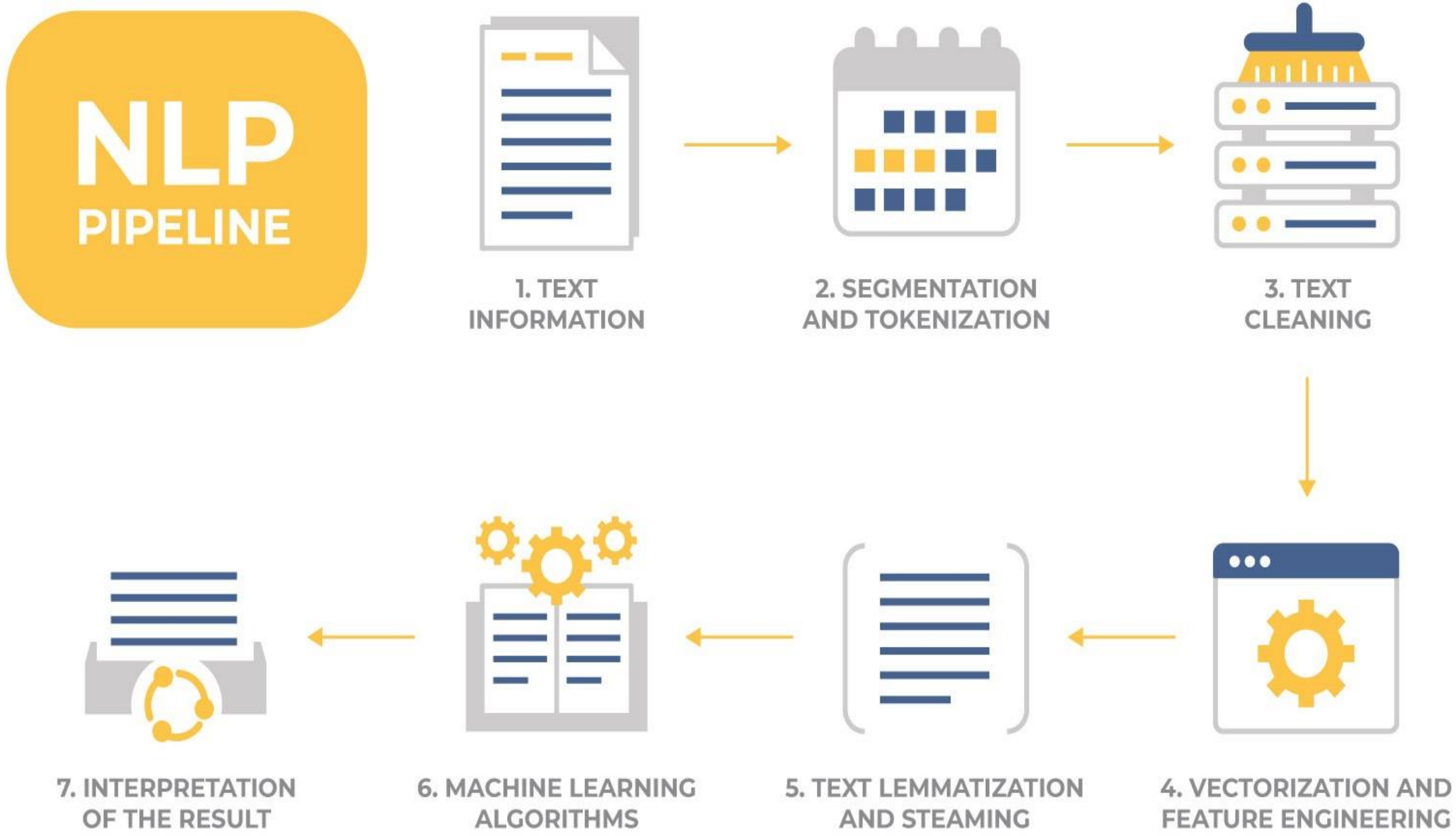
Step 5: RNN Model
Train an RNN model to learn the mapping between English and French sequences. The RNN processes one word at a time, maintaining a hidden state that captures the context of the input sequence.

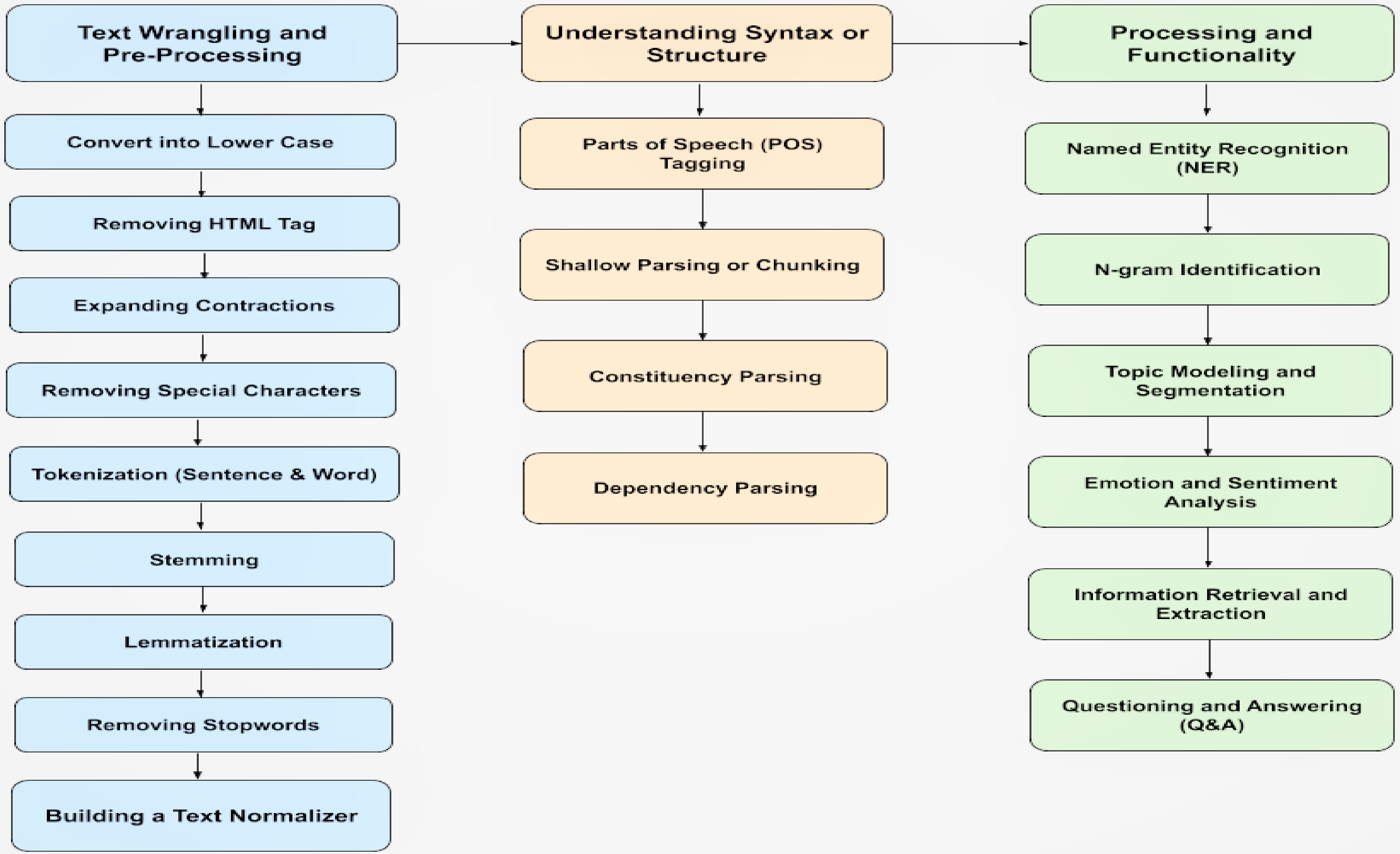
Step 6: Translation
Now, given a new English sentence, the trained RNN can generate a sequence of French words as output.
- Input (English): "I enjoy deep learning."
- Output (French): "J'apprécie l'apprentissage profond."

In this example, the RNN has learned the patterns and relationships between English and French words during training and is able to generate reasonable translations for new input sequences.

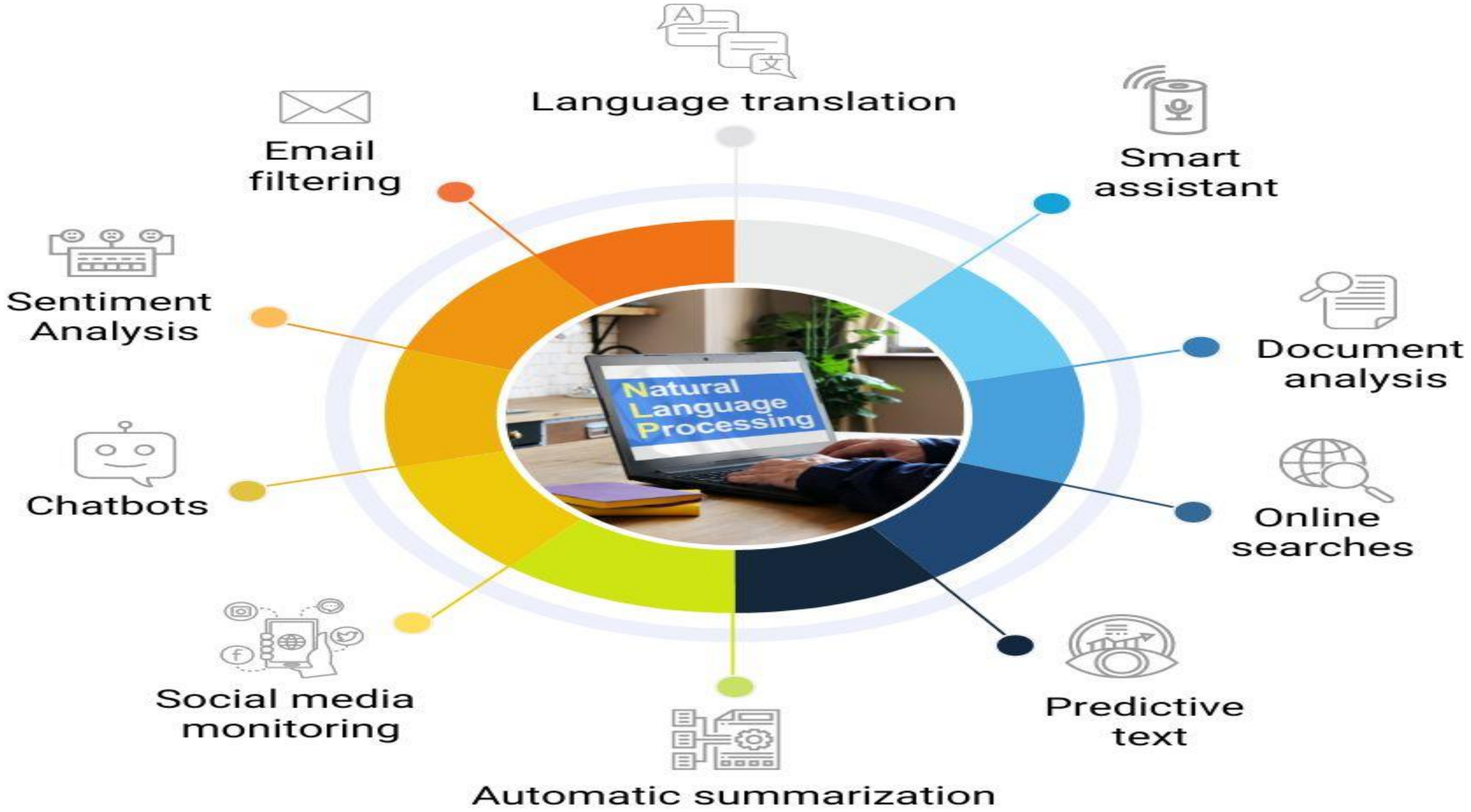

Text
Video
Audio
NLP
Relationship Extraction
Entity Extraction
Topic Classification
Sentiment Analysis
Raw Language Processing

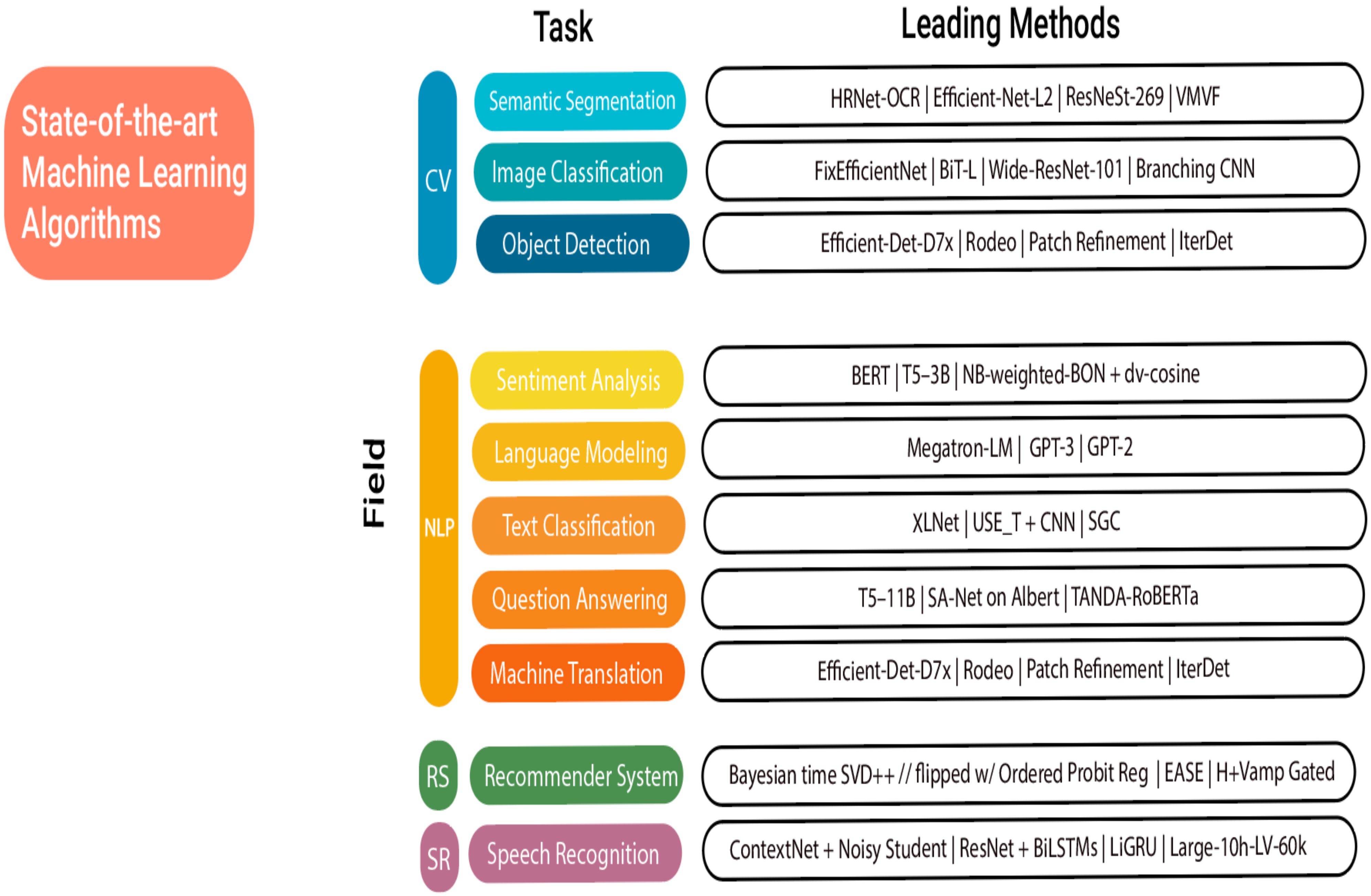
Standard NLP Workflow


NLP
PIPELINE
1. TEXT INFORMATION
2. SEGMENTATION AND TOKENIZATION
3. TEXT CLEANING
4. VECTORIZATION AND FEATURE ENGINEERING
5. TEXT LEMMATIZATION AND STEAMING
6. MACHINE LEARNING ALGORITHMS
7. INTERPRETATION OF THE RESULT


Text Wrangling and Pre-Processing
Convert into Lower Case
Removing HTML Tag
Expanding Contractions
Removing Special Characters
Tokenization (Sentence & Word)
Stemming
Lemmatization
Removing Stopwords
Building a Text Normalizer
Understanding Syntax or Structure
Parts of Speech (POS) Tagging
Shallow Parsing or Chunking
Constituency Parsing
Dependency Parsing
Processing and Functionality
Named Entity Recognition (NER)
N-gram Identification
Topic Modeling and Segmentation
Emotion and Sentiment Analysis
Information Retrieval and Extraction
Questioning and Answering (Q&A)

Applications of Natural Language Processing




State-of-the-art Machine Learning Algorithms
Task
Leading Methods
Field
CV
Semantic Segmentation
HRNet-OCR | Efficient-Net-L2 | ResNeSt-269 | VMVF
Image Classification
FixEfficientNet | BiT-L | Wide-ResNet-101 | Branching CNN
Object Detection
Efficient-Det-D7x | Rodeo | Patch Refinement | IterDet
NLP
Sentiment Analysis
BERT | T5–3B | NB-weighted-BON + dv-cosine
Language Modeling
Megatron-LM | GPT-3 | GPT-2
Text Classification
XLNet | USE_T + CNN | SGC
Question Answering
T5–11B | SA-Net on Albert | TANDA-RoBERTa
Machine Translation
Efficient-Det-D7x | Rodeo | Patch Refinement | IterDet
RS
Recommender System
Bayesian time SVD++ // flipped w/ Ordered Probit Reg | EASE | H+Vamp Gated
SR
Speech Recognition
ContextNet + Noisy Student | ResNet + BiLSTMs | LiGRU | Large-10h-LV-60k

Applications

NLP

- Chatbots
- Neural Machine Translation
- Text Summarization
- Question Answering
- Language Modelling
- Autocomplete
- Autocorrect

Computer Vision

- Agriculture
- Segmentation
- Post Estimation
- Medical Diagnosis
- Deep Fakes
- Retail
- Financial Services

QUESTIONS

THANK YOU