Roadmap To Learn Machine Learning

Mathematics

- Probability • Discrete
- Statistics

Programming

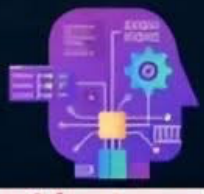
- Python • R

Database

- MySQL • MongoDB

Machine Learning

- Scikit learn Supervised learning
- Unsupervised learning
- Reinforcement learning



Machine Learning

ML Libraries and Non- ML Libraries

ML Algorithms

- Linear Logistic Regression
- KNN
- K-means
- Random forest & more!

Deep Learning

- TensorFlow
- Keras
- Neural Networks
- CNN, RNN, GAN, LSTM

Data Visualization Tools

- Tableau
- Qlikview
- PowerBI

ML Engineer

12 Must Know GenAI Terms

Created by Brij Kishore Pandey

LLM

Large Language Model

Advanced AI systems trained on vast text datasets to understand and generate human-like text, serving as the foundation for modern conversational AI and content generation.

Transformers

Transformer Architecture

Revolutionary neural network architecture using self-attention mechanisms to process sequential data, enabling breakthrough capabilities in language understanding and generation.

Prompt Engineering

AI Instruction Design

Strategic formulation of inputs to AI models to achieve desired outputs, combining precise instructions, context, and constraints for optimal results.

Fine-tuning

Model Specialization

Process of adapting pre-trained AI models to specific tasks or domains by training on specialized datasets, enhancing performance for targeted applications.

Embeddings

Vector Representations

Numerical representations of text, images, or data in high-dimensional space, enabling semantic search, similarity comparisons, and efficient AI processing.

RAG

Retrieval Augmented Generation

Technique combining knowledge retrieval with text generation to produce accurate, factual responses by accessing external information sources during generation.

Tokens

Text Units

01
10

Fundamental units of text processing in AI models, representing words, subwords, or characters, determining model capacity and processing limitations.

Hallucination

AI Fabrication

Phenomenon where AI models generate plausible but factually incorrect information, a key challenge in ensuring reliable AI outputs.

Zero-shot

Zero-shot Learning

AI capability to perform tasks without specific training examples, using general knowledge to understand and execute new instructions.

Chain-of-Thought

Reasoning Process

Prompting technique encouraging AI models to break down complex problems into step-by-step reasoning, improving accuracy and explainability.

Context Window

Input Capacity

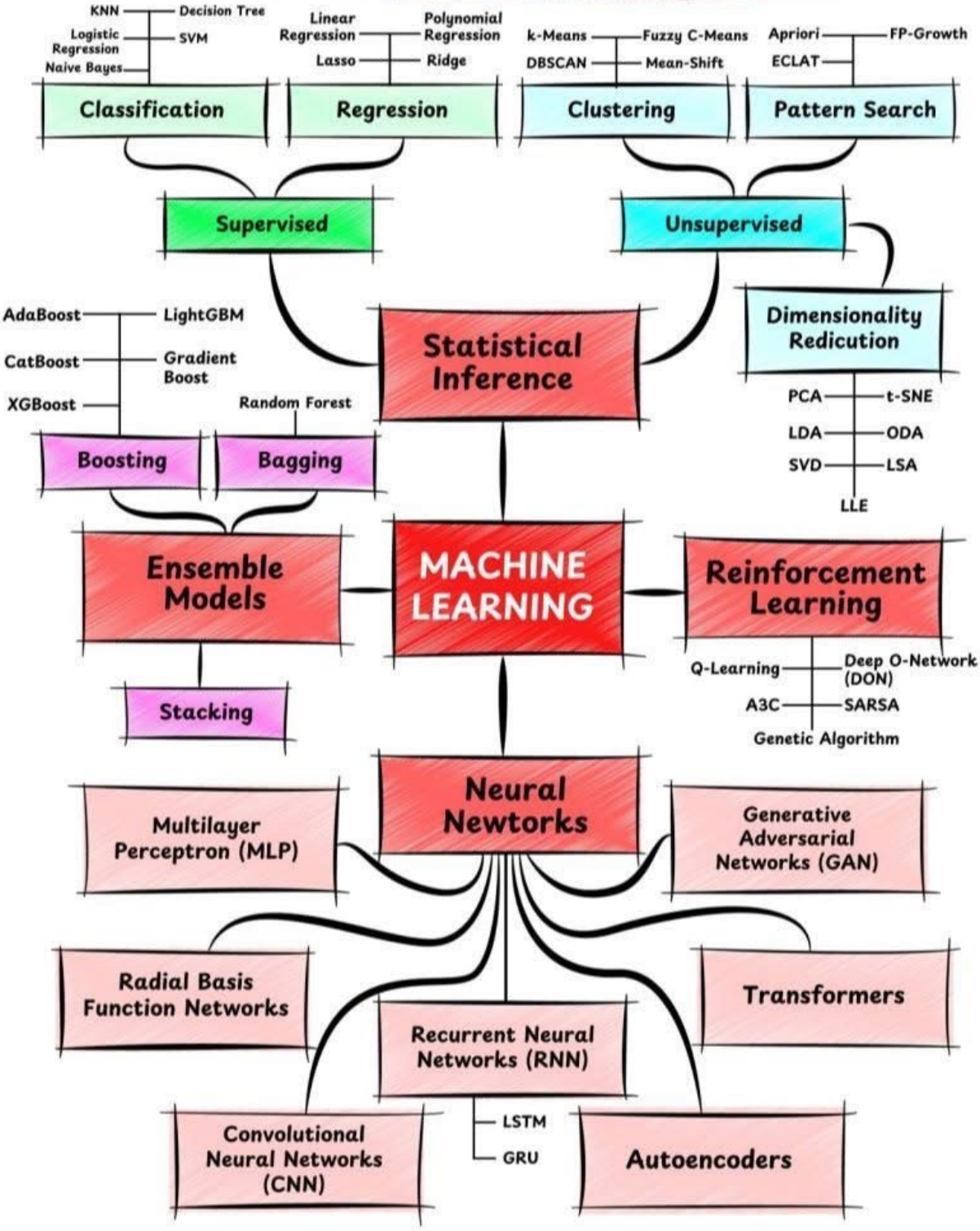
Maximum amount of text an AI model can process in a single interaction, affecting its ability to maintain coherence and reference information.

Temperature

Randomness Parameter

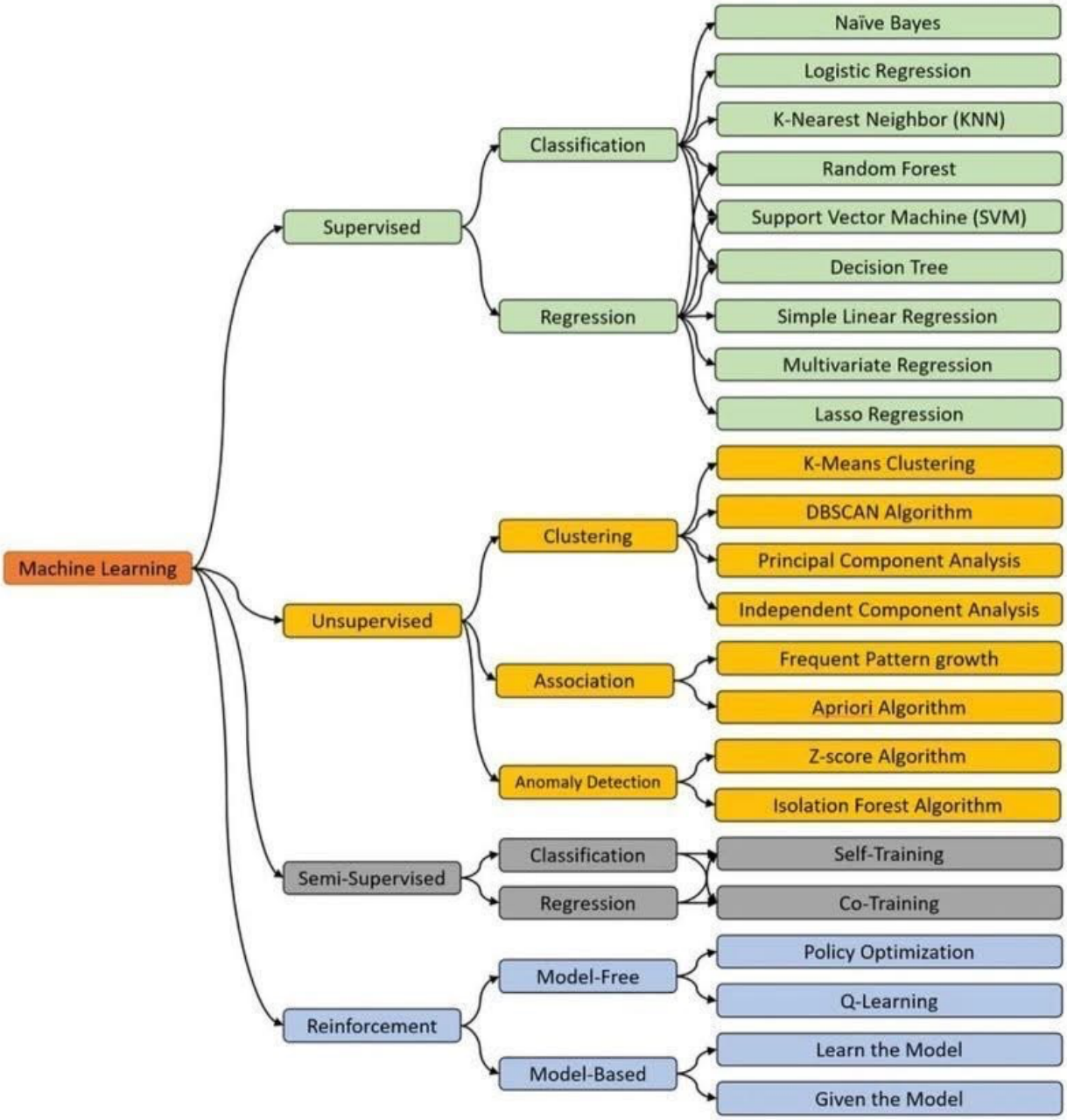
Control parameter affecting the randomness and creativity of AI outputs, balancing between deterministic responses and creative variations.

Popular Models For Machine Learning



Machine Learning Algorithms

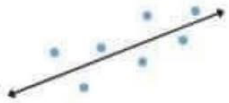
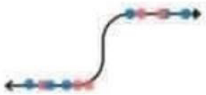
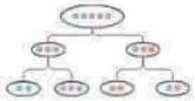
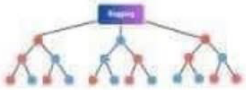
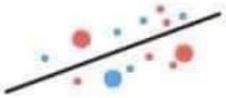
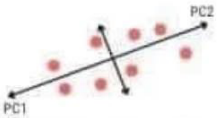
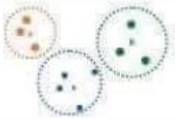
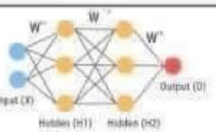
(Every data scientist must know)



Machine Learning Time Complexity

DataInterview.com

n: data size
p: Number of features
T: Number of trees
l: Number of iterations
m: Number of components
h: Number of hidden units
k: number of clusters

ML Algorithms	Training Time	Inference Time
Linear Regression 	$O(np^2 + p^3)$	$O(p)$
Logistic Regression 	$O(np^2 + p^3)$	$O(p)$
Naive Bayes 	$O(np)$	$O(p)$
Decision Tree 	Avg: $O(T \cdot n \log n)$ Worst: $O(n^2)$	Avg: $O(T \cdot n \log n)$ Worst: $O(n)$
Random Forest 	$O(T \cdot n \log n)$	$O(T \cdot \log n)$
Gradient Boosted Trees 	$O(T \cdot n \log n)$	$O(T \cdot \log n)$
Principal Component 	$O(np^2 + p^3)$	$O(pm)$
K-Nearest Neighbor 	$O(1)$	$O(np)$
K-Means 	$O(l \cdot k \cdot n \cdot p)$	$O(k \cdot p)$
Dense Neural Networks 	$O(l \cdot n \cdot p \cdot h)$	$O(p \cdot h)$