AIML-501 PORTFOLIO ARTIFACT · WORKSHOP 1

Machine Learning Algorithm Types: A Visual Framework

Fourteen widely used algorithms, sorted by how they learn, the kind of data they work on, and what people use them for in the real world. The lower panels show how the algorithms connect and where to start.

Bindu Bhargava Reddy Chintam
M.S. student, Indiana Wesleyan University
AIML-501: Model Development
October 2026

HOW TO READ THIS FRAMEWORK

- **Supervised.** Learns from examples that already carry the right answer (a label).
- **Unsupervised.** Gets no labels and finds structure, such as groups or main directions, by itself.
- **Unsupervised: self-supervised and generative.** No human labels; the data itself supplies the target, such as the next word or a hidden word, or the model learns to create new data.

APPLICATION DOMAINS

Tabular data · Computer vision · NLP · Generative AI
Dashed = also used

WHERE EACH ALGORITHM SITS IN THE AI FAMILY

ARTIFICIAL INTELLIGENCE
Any system that performs tasks that need human-like intelligence

MACHINE LEARNING
Learns patterns from data instead of hand-written rules.
Classic ML here: Linear and Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, SVM, KNN, K-means, PCA

DEEP LEARNING
Neural networks with many layers.
Here: CNN, BERT

GENERATIVE AI
Deep models that create new text, images, or audio.
Here: GPT, GAN, Diffusion models

SUPERVISED — Learns from labeled examples

Eight algorithms that predict a number or a category

01 Linear Regression — SUPERVISED
Linear model
Tabular

HOW IT WORKS
Draws the straight line (or flat surface) that best fits past data points, then reads new predictions off that line.

REAL-WORLD USE
Forecasting monthly sales; estimating a house price from its size and location.

02 Logistic Regression — SUPERVISED
Linear model
Tabular · NLP

HOW IT WORKS
Turns a weighted sum of the inputs into a probability between 0 and 1, then picks a class when it passes a cutoff.

REAL-WORLD USE
Credit default risk scores; simple spam filters built on word counts.

03 Decision Tree — SUPERVISED
Tree model
Tabular

HOW IT WORKS
Asks a chain of yes or no questions about the features and follows the answers down to a label. Easy to read and explain.

REAL-WORLD USE
Loan approval rules; sorting fruit photos by color share (my Workshop 1.4 project).

04 Random Forest — SUPERVISED
Tree ensemble (bagging)
Tabular

HOW IT WORKS
Grows hundreds of trees, each on a random slice of rows and features, and lets them vote. The vote cancels out single-tree mistakes.

REAL-WORLD USE
Credit card fraud detection; ranking which genes matter most in a disease study.

05 Gradient Boosting — SUPERVISED
Tree ensemble (boosting): XGBoost, LightGBM
Tabular

HOW IT WORKS
Builds small trees one after another. Each new tree focuses on the errors the earlier trees left behind.

REAL-WORLD USE
Ad click and churn prediction; a frequent winner of tabular data competitions.

06 Support Vector Machine — SUPERVISED
Margin-based model
Tabular · NLP · Vision

HOW IT WORKS
Finds the boundary with the widest gap between classes. A kernel lets that boundary curve when a straight line will not work.

REAL-WORLD USE
Sorting news articles by topic; recognizing handwritten digits on small datasets.

07 K-Nearest Neighbors — SUPERVISED
Distance-based model
Tabular · Vision

HOW IT WORKS
Keeps every training example and labels a new one by a majority vote of its k closest neighbors. No real training step.

REAL-WORLD USE
"Customers like you also bought" suggestions; finding look-alike images.

08 Convolutional Neural Network — SUPERVISED
Deep learning (CNN)
Vision · GenAI

HOW IT WORKS
Slides small filters across an image to spot edges, then shapes, then whole objects, one layer at a time.

REAL-WORLD USE
Face unlock on phones; flagging chest X-rays for review; reading road signs.

UNSUPERVISED — Finds structure without labels

Classic methods on the left; self-supervised and generative deep models on the right

CLASSIC UNSUPERVISED

09 K-means Clustering — UNSUPERVISED
Clustering, distance-based
Tabular · Vision

HOW IT WORKS
Places k center points, assigns each item to the nearest center, moves each center to the middle of its group, and repeats until nothing changes.

REAL-WORLD USE
Customer segmentation for marketing; shrinking a photo down to a few colors.

10 Principal Component Analysis — UNSUPERVISED
Dimensionality reduction (PCA)
Tabular · Vision

HOW IT WORKS
Rotates the data to find the directions with the most variation and keeps only the top few, so many columns become a handful.

REAL-WORLD USE
Turning thousands of gene measurements into a 2-D plot; speeding up other models.

SELF-SUPERVISED AND GENERATIVE (DEEP LEARNING)

11 BERT — SELF-SUPERVISED
Transformer encoder
NLP

HOW IT WORKS
Reads a sentence in both directions at once. It learns by guessing words that were hidden, then gets fine-tuned with labels for a task.

REAL-WORLD USE
Helping search engines understand queries; sentiment analysis of reviews.

12 GPT — SELF-SUPERVISED
Transformer decoder (LLM)
NLP · GenAI

HOW IT WORKS
Predicts the next word piece again and again, writing text one token at a time. Later tuning with human feedback shapes its answers.

REAL-WORLD USE
Chat assistants such as ChatGPT; drafting code, emails, and summaries.

13 Generative Adversarial Network — GENERATIVE
Generative model (GAN)
Vision · GenAI

HOW IT WORKS
A generator makes fake samples while a discriminator tries to catch them. Both improve until the fakes look real.

REAL-WORLD USE
Making synthetic training images; sharpening low-resolution photos.

14 Diffusion Model — GENERATIVE
Generative model (e.g., Stable Diffusion)
Vision · GenAI

HOW IT WORKS
Learns to remove noise from images step by step, so it can start from pure noise and build a brand-new picture that matches a text prompt.

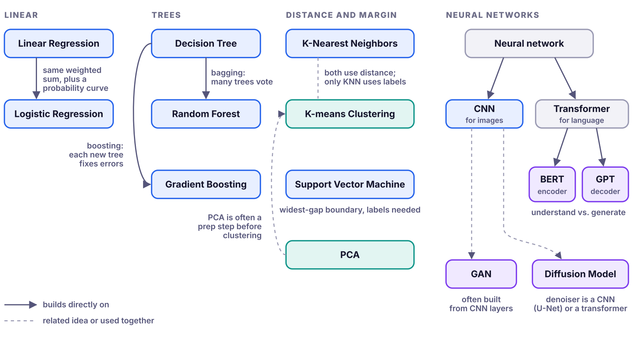
REAL-WORLD USE
Text-to-image tools; quick concept art and product mockups.

DOMAIN MAP: WHERE EACH ALGORITHM IS USED

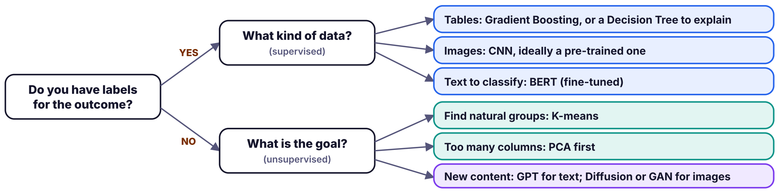
ALGORITHM	TABULAR	VISION	NLP	GENAI
01 Linear Regression	●			
02 Logistic Regression	●		○	
03 Decision Tree	●			
04 Random Forest	●			
05 Gradient Boosting	●			
06 Support Vector Machine	●	○	○	
07 K-Nearest Neighbors	●	○		
08 CNN		●		○
09 K-means Clustering	●	○		
10 PCA	●	○		
11 BERT			●	
12 GPT			●	●
13 GAN		●		●
14 Diffusion Model		●		●

Filled dot = main domain. Ring = also used there. The pattern is clear: classic ML owns tabular data, CNNs own vision, transformers own language, and generative AI sits where deep vision and language models meet.

FAMILY MAP: HOW THE ALGORITHMS BUILD ON EACH OTHER



QUICK GUIDE: PICKING A FIRST ALGORITHM TO TRY



Why this order works

1. Labels decide the learning style first.
2. The data type then narrows the family: trees for tables, CNNs for images, transformers for language.
3. Start simple, measure, then move up.

A clear tree beats a black box that no one can check, unless accuracy truly needs it.

Key sources: Breiman (2001); Breiman et al. (1984); Chen & Guestrin (2016); Cortes & Vapnik (1995); Cover & Hart (1967); Cox (1958); Devlin et al. (2019); Friedman (2001); Goodfellow et al. (2014); Ho et al. (2020); LeCun et al. (1998); MacQueen (1967); Pearson (1901); Radford et al. (2018); Rombach et al. (2022); Ronneberger et al. (2015); Trojahn & Costa (2024); Vaswani et al. (2017). Full APA 7 list on the portfolio page and in the accompanying document.

Designed by Bindu Bhargava Reddy Chintam
bhargavchintam.github.io/Portfolio