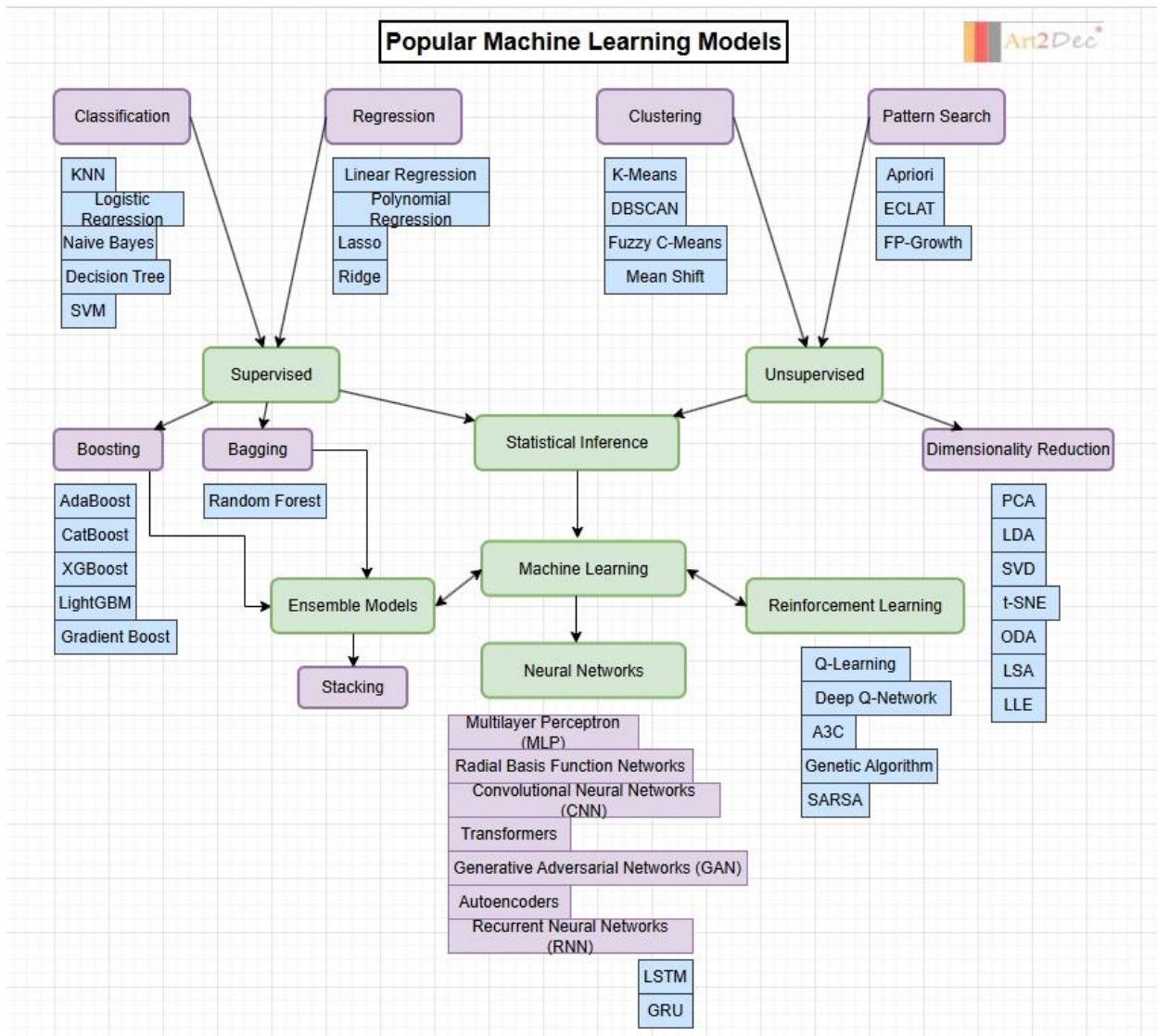


# Machine learning taxonomy diagram.



Here is a short description of this machine learning taxonomy diagram.

## Overall Structure

This diagram presents a hierarchical taxonomy of popular machine learning models, organized into main categories and subcategories with clear relationships between different approaches.

## Main Categories (Top Level)

The diagram divides machine learning into four primary task types:

1. **Classification** - Models that assign discrete labels to inputs
2. **Regression** - Models that predict continuous numerical values
3. **Clustering** - Models that group similar data points without labels
4. **Pattern Search** - Models that discover patterns and associations in data

## Primary Learning Paradigms (Second Level)

These task types feed into two fundamental learning approaches:

### Supervised Learning

Learns from labeled training data where inputs are mapped to known outputs. This paradigm includes:

#### Classification Models:

- **KNN** (K-Nearest Neighbors) - Instance-based learning using proximity
- **Logistic Regression** - Probabilistic classification using logistic function
- **Naive Bayes** - Probabilistic classifier based on Bayes' theorem
- **Decision Tree** - Tree-structured decision rules
- **SVM** (Support Vector Machine) - Finds optimal separating hyperplane

#### Regression Models:

- **Linear Regression** - Models linear relationships between variables
- **Polynomial Regression** - Extends linear regression with polynomial terms
- **Lasso** - Linear regression with L1 regularization
- **Ridge** - Linear regression with L2 regularization

### Unsupervised Learning

Learns patterns from unlabeled data without predefined outputs. This paradigm includes:

#### Clustering Models:

- **K-Means** - Partitions data into K clusters using centroids
- **DBSCAN** (Density-Based Spatial Clustering) - Finds clusters based on density

- **Fuzzy C-Means** - Soft clustering where points can belong to multiple clusters
- **Mean Shift** - Non-parametric clustering using density estimation

### **Pattern Search Models:**

- **Apriori** - Association rule mining algorithm
- **ECLAT** (Equivalence Class Transformation) - Vertical data format for mining
- **FP-Growth** (Frequent Pattern Growth) - Efficient pattern mining without candidate generation

## **Advanced Techniques and Meta-Learning**

### **Ensemble Methods**

Techniques that combine multiple models to improve performance:

#### **Boosting (Sequential ensemble):**

- **AdaBoost** - Adaptive boosting, focuses on misclassified examples
- **CatBoost** - Gradient boosting for categorical features
- **XGBoost** - Extreme gradient boosting with regularization
- **LightGBM** - Gradient boosting using histogram-based algorithms
- **Gradient Boost** - General gradient boosting framework

#### **Bagging (Parallel ensemble):**

- **Random Forest** - Ensemble of decision trees with random feature selection

#### **Stacking:**

- Combines predictions from multiple models using a meta-learner
- Feeds into **Ensemble Models** node

## **Statistical Inference**

A bridging concept connecting supervised and unsupervised learning, representing the theoretical foundation for probabilistic and statistical approaches in machine learning.

## **Core Machine Learning Hub**

The central **Machine Learning** node connects to four major branches:

### **1. Neural Networks**

Deep learning approaches using interconnected layers of artificial neurons:

## Architectures:

- **Multilayer Perceptron (MLP)** - Fully connected feedforward networks
- **Radial Basis Function Networks** - Uses radial basis functions as activation
- **Convolutional Neural Networks (CNN)** - Specialized for spatial data (images)
- **Transformers** - Attention-based architecture for sequence modeling
- **Generative Adversarial Networks (GAN)** - Two networks competing (generator vs discriminator)
- **Autoencoders** - Unsupervised learning for dimensionality reduction and feature learning
- **Recurrent Neural Networks (RNN)** - Processes sequential data with memory
  - **LSTM** (Long Short-Term Memory) - RNN variant handling long-term dependencies
  - **GRU** (Gated Recurrent Unit) - Simplified LSTM variant

## 2. Reinforcement Learning

Learning through interaction with an environment to maximize rewards:

- **Q-Learning** - Value-based learning using Q-tables
- **Deep Q-Network** - Q-learning with deep neural networks
- **A3C** (Asynchronous Advantage Actor-Critic) - Parallel actor-critic approach
- **Genetic Algorithm** - Evolution-inspired optimization
- **SARSA** (State-Action-Reward-State-Action) - On-policy temporal difference learning

## 3. Ensemble Models

Central hub for combining multiple learning algorithms (connects to Boosting, Bagging, and Stacking methods described above)

## 4. Dimensionality Reduction

Techniques for reducing the number of features while preserving important information:

- **PCA** (Principal Component Analysis) - Linear transformation to principal components
- **LDA** (Linear Discriminant Analysis) - Supervised dimensionality reduction maximizing class separability
- **SVD** (Singular Value Decomposition) - Matrix factorization technique
- **t-SNE** (t-Distributed Stochastic Neighbor Embedding) - Non-linear visualization technique
- **ODA** (likely Optimal Discriminant Analysis)
- **LSA** (Latent Semantic Analysis) - Dimensionality reduction for text data
- **LLE** (Locally Linear Embedding) - Non-linear manifold learning

## Key Relationships and Connections

1. **Task-to-Paradigm Flow:** Classification and Regression feed into Supervised Learning; Clustering feeds into Unsupervised Learning; Pattern Search also connects to Unsupervised Learning
2. **Statistical Inference Bridge:** Acts as a bidirectional connection between Supervised and Unsupervised paradigms, highlighting their shared statistical foundations
3. **Ensemble-to-Core:** Both Boosting and Bagging methods converge into Ensemble Models, which connects to the central Machine Learning hub
4. **Neural Network Hierarchy:** Shows evolution from simple MLPs to complex architectures like Transformers and GANs, with RNNs branching into LSTM and GRU variants
5. **Holistic Integration:** All pathways ultimately connect to the central "Machine Learning" concept, demonstrating that these diverse approaches are all part of a unified field

This taxonomy provides a comprehensive map of machine learning approaches, showing how different techniques relate to each other and serve different purposes within the broader ML ecosystem.

Contacts: [igor.lukyanov@appservgrid.com](mailto:igor.lukyanov@appservgrid.com) Art2Dec Soft Lab.

11.14.2025

---