Machine Learning Taxonomy with Summaries

Machine Learning
--- 1) Learning Paradigms
  --- Supervised Learning: Learn from labeled data to predict outputs.
    --- Regression: Predict continuous numeric values.
      --- Linear Regression: Fit a straight line to data points.
      --- Ridge/Lasso Regression: Linear regression with regularization.
    --- Classification: Predict discrete class labels.
      --- Logistic Regression: Model probability of classes using sigmoid.
      --- Decision Trees (CART): Tree-based models splitting data by features.
      --- Support Vector Machines: Maximize margin between classes with a hyperplane.
  --- Unsupervised Learning: Discover patterns without labeled outputs.
    --- Clustering: Group similar data points.
      --- k-Means: Partition data into k clusters minimizing within-cluster variance.
      --- DBSCAN: Density-based clustering of arbitrary-shaped clusters.
    --- Dimensionality Reduction: Reduce feature space while preserving structure.
      --- PCA: Linear technique for maximizing variance directions.
      --- t-SNE/UMAP: Nonlinear techniques for visualizing high-dimensional data.
  --- Semi-Supervised Learning: Use small labeled + large unlabeled data to train.
  --- Self-Supervised Learning: Create pseudo-labels from data structure itself.
  --- Reinforcement Learning: Learn optimal actions by trial and error maximizing rewards.
    --- Value-based (e.g., Q-Learning): Learn expected rewards of actions.
    --- Policy-based (e.g., Policy Gradients): Learn directly the action probabilities.

--- 2) Model Types
  --- Parametric Models: Fixed number of parameters regardless of data size.
    --- Linear Models: Predict outputs with weighted sums of inputs.
    --- Neural Networks (fixed architecture): Layers of neurons with trainable weights.
  --- Non-Parametric Models: Complexity grows with data size.
    --- k-Nearest Neighbors: Predict based on closest training points.
    --- Decision Trees: Split data hierarchically without fixed parameters.

--- 3) Algorithms / Approaches
  --- Linear Models: Use weighted sums of inputs to predict outputs.
    --- Linear/Logistic Regression: Predict continuous values or probabilities.
    --- Perceptron Algorithm: Early linear classifier updating weights on mistakes.
  --- Decision Trees: Learn tree-like rules for prediction.
    --- CART, ID3, C4.5: Different algorithms for building decision trees.
  --- Ensemble Methods: Combine multiple models for better performance.
    --- Bagging (Random Forests): Aggregate bootstrapped trees to reduce variance.
    --- Boosting (AdaBoost, XGBoost): Sequentially add models to correct errors.
  --- Support Vector Machines: Classifiers maximizing margins with support vectors.
  --- k-Nearest Neighbors: Predict by majority vote or average of nearest data points.
  --- Neural Networks: Layers of connected neurons learning complex patterns.
    --- Feedforward Networks (MLP): Basic fully-connected layers for regression/classification.
    --- Convolutional Neural Networks (CNNs): Specialized for images and spatial data.
    --- Recurrent Neural Networks (RNNs): Handle sequential/time-series data.
    --- Generative Neural Networks: Learn data distributions to generate new samples.
      --- GANs: Compete generator and discriminator networks to synthesize realistic data.
      --- VAEs: Encode data into probabilistic latent space and decode samples.
    --- Recent Architecture (e.g., Transformers): Attention-based networks excelling in NLP and vision.

- Probabilistic Models: Use probability distributions for predictions.
  - Naive Bayes: Assumes feature independence for fast probabilistic classification.
  - Hidden Markov Models (HMMs): Model sequences with hidden states emitting observations.
- Clustering Algorithms: Group unlabeled data based on similarity.
  - k-Means: Minimize intra-cluster distances.
  - Hierarchical Clustering: Build nested clusters in a tree-like structure.

- 4) Tasks / Applications
  - Computer Vision: Understand and analyze images and videos.
    - Image Classification: Assign labels to entire images.
    - Object Detection: Locate and classify objects within images.
    - Image Segmentation: Partition images into labeled regions.
  - Natural Language Processing (NLP): Analyze and generate human language.
    - Text Classification: Assign categories to text (e.g., spam detection).
    - Machine Translation: Translate text between languages.
    - Named Entity Recognition: Identify proper names in text.
  - Speech Recognition: Convert spoken language to text.
    - Speech-to-Text: Transcribe audio recordings into written words.
  - Recommendation Systems: Suggest items based on user preferences.
    - Collaborative Filtering: Predict interests by similarities between users/items.
    - Content-Based Filtering: Recommend based on item features.
  - Time Series Forecasting: Predict future values from historical sequential data.
    - ARIMA, LSTM Models: Capture temporal patterns for forecasting.
  - Anomaly Detection: Identify unusual patterns deviating from normal behavior.
    - One-Class SVM, Isolation Forest: Algorithms for outlier detection.

- 5) Evaluation Metrics
  - Classification Metrics: Measure classification performance.
    - Accuracy: Proportion of correct predictions.
    - Precision, Recall, F1-Score: Trade-offs between false positives/negatives.
    - ROC AUC: Measure of classifier separability.
    - Log Loss: Penalizes confident wrong predictions.
  - Regression Metrics: Measure numeric prediction performance.
    - Mean Squared Error (MSE): Average squared prediction error.
    - Mean Absolute Error (MAE): Average absolute prediction error.
    - R-squared (R^2): Proportion of variance explained by the model.