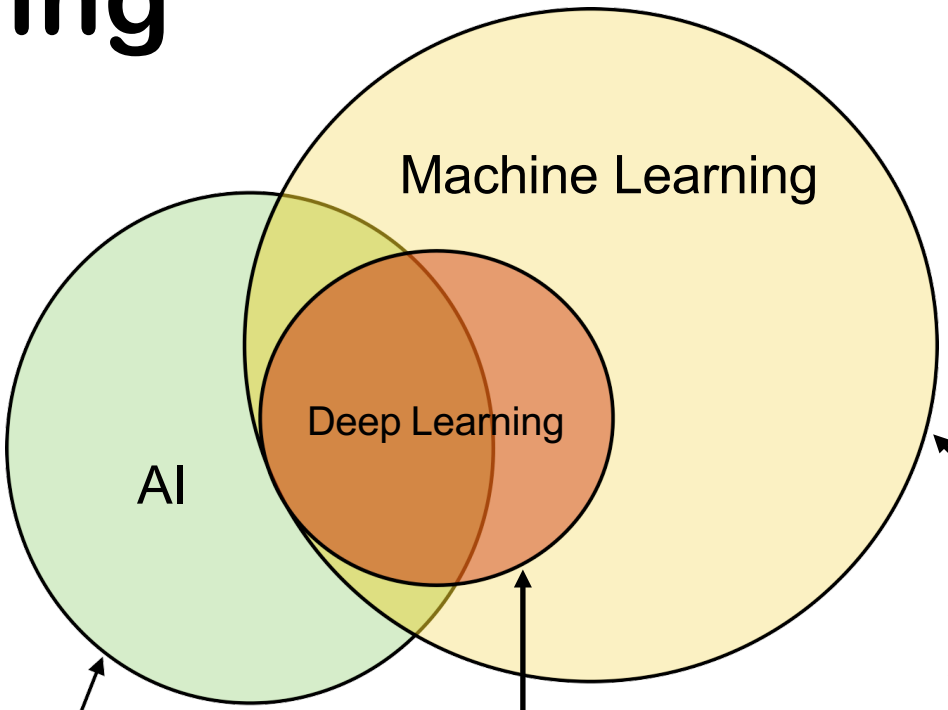
Machine Learning
CSCE 5215

Machine Learning Terminologies

Instructor: Zeenat Tariq

Machine Learning, AI, and Deep Learning



Algorithms that learn models/representations/ rules automatically from data/examples

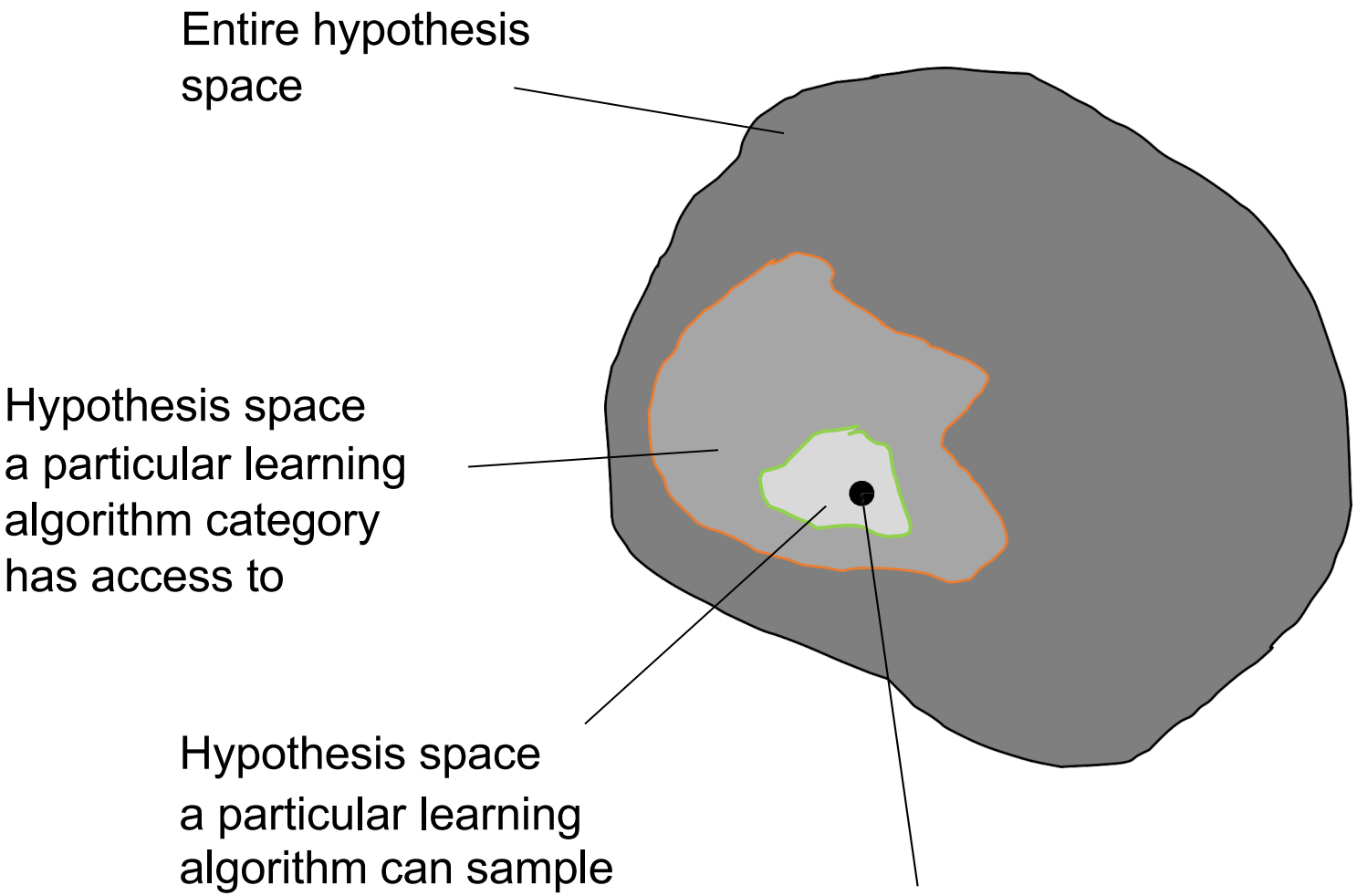
A non-biological system that is intelligent through rules

Algorithms that parameterize multilayer neural networks that then learn representations of data with multiple layers of abstraction

ML Terminology (Part 1)

- **Training example**: A row in the table representing the dataset. Synonymous to an observation, training record, training instance, training sample (in some contexts, sample refers to a collection of training examples)
- **Feature**: a column in the table representing the dataset. Synonymous to predictor, variable, input, attribute, covariate
- **Targets**: What we want to predict. Synonymous to outcome, output, ground truth, response variable, dependent variable, (class) label (in classification)
- **Output / prediction**: use this to distinguish from targets; here, means output from the model

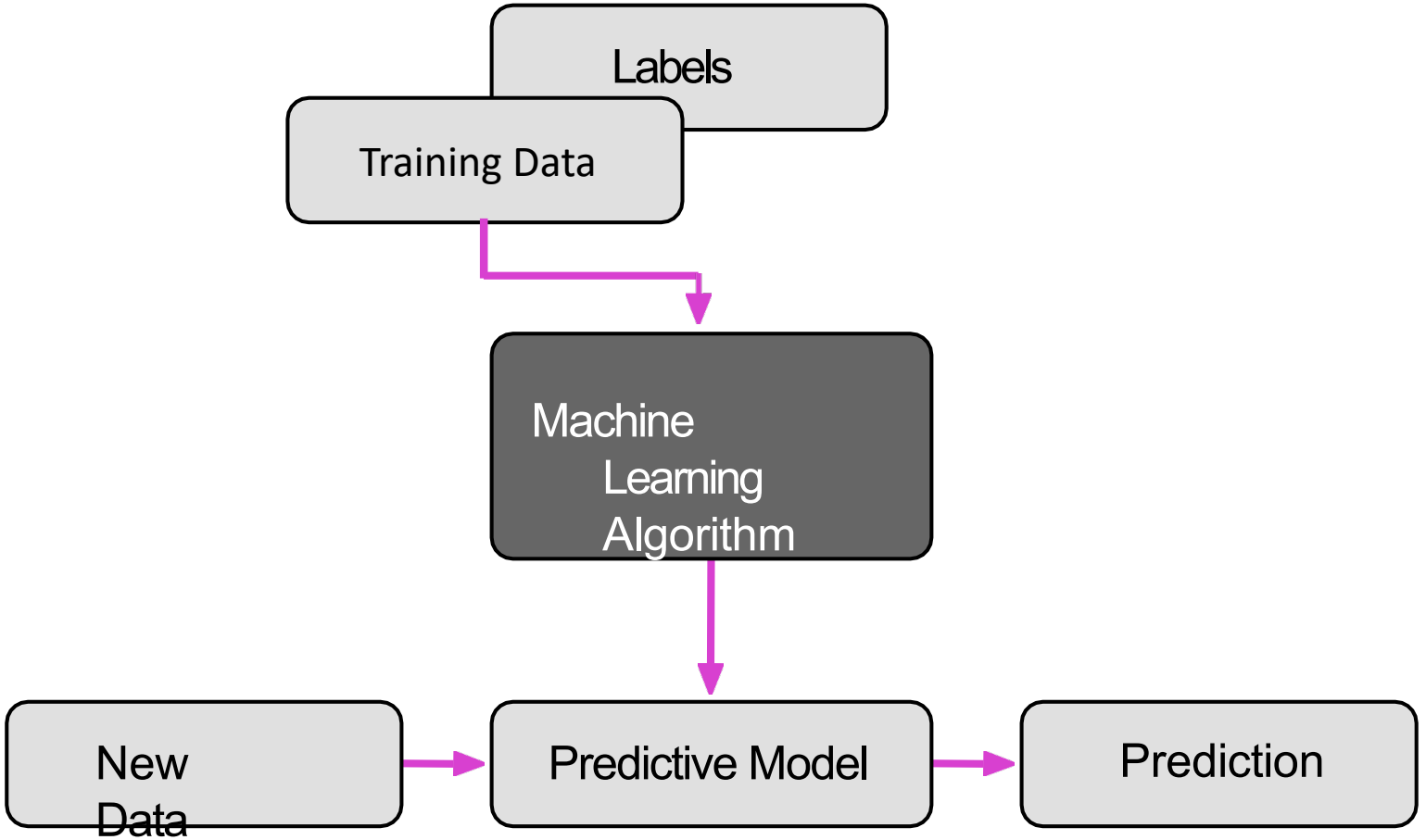
Hypothesis Space

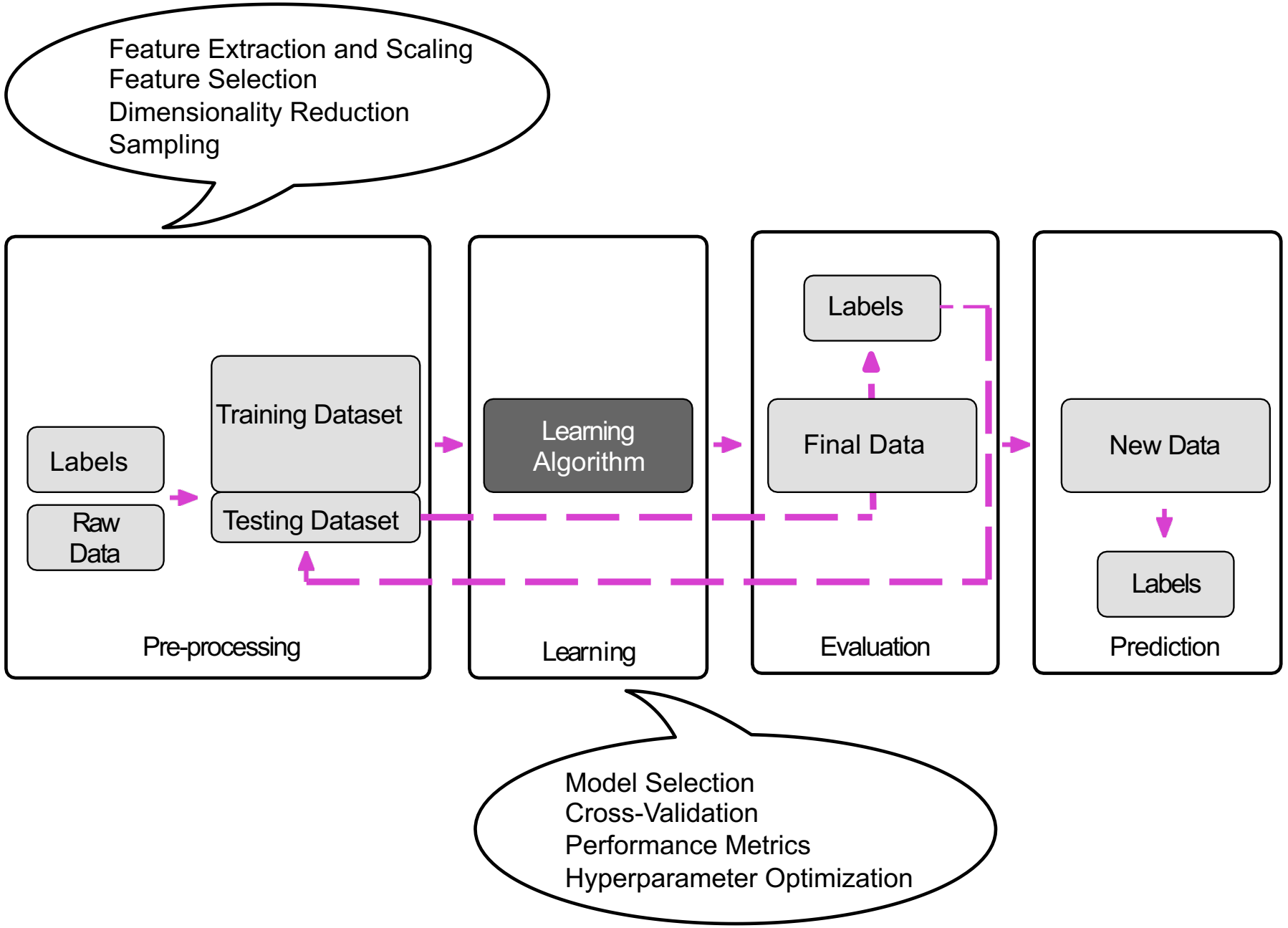


Classes of Machine Learning Algorithms

- Generalized linear models (e.g., Logistic Regression)
- Support vector machines (e.g., Linear SVM, Radius base function Kernel SVM)
- Artificial neural networks (e.g., Multilayer perceptron CNN, RNN)
- Tree- or rule-based models (e.g., Decision Tree)
- Graphical models (e.g., Bayesian Networks)
- Ensembles (e.g., Random Forest)
- Instance-based learners (e.g., nearest neighbors)

Supervised Learning Workflow -- Overview




Feature Extraction and Scaling
Feature Selection
Dimensionality Reduction
Sampling
Labels
Raw Data
Training Dataset
Testing Dataset
Pre-processing
Learning Algorithm
Learning
Labels
Final Data
Evaluation
New Data
Labels
Prediction
Model Selection
Cross-Validation
Performance Metrics
Hyperparameter Optimization

Steps for approaching a Machine Learning Application

1. Define the problem to be solved.
2. Collect (labeled) data.
3. Choose an algorithm class.
4. Choose an optimization metric or measure for learning the model.
5. Choose a metric or measure for evaluating the model.

Objective Functions

- Maximize the posterior probabilities (e.g., naive Bayes)
- Maximize a fitness function (genetic programming)
- Maximize the total reward/value function (reinforcement learning)
- Maximize information gain/minimize child node impurities (CART decision tree classification)
- Minimize a mean squared error cost (or loss) function (CART, decision tree regression, linear regression, adaptive linear neurons, ...)
- Maximize log-likelihood or minimize cross-entropy loss (or cost) function
- Minimize hinge loss (support vector machine)

Optimization Methods for Different Learning Algorithms

- Combinatorial search, greedy search (e.g., decision trees)
- Unconstrained convex optimization (e.g., Linear regression, logistic regression)
- Constrained convex optimization (e.g., SVM)
- Nonconvex optimization, here: using backpropagation, chain rule, reverse autodiff (e.g., neural networks)
- Constrained nonconvex optimization (e.g., semi-adversarial network)

Evaluation -- Misclassification Error

$$L(\hat{y}, y) = \begin{cases} 0 & \textbf{if } \hat{y} = y \\ 1 & \textbf{if } \hat{y} \neq y \end{cases}$$

$$ERR_{\mathcal{D}\textbf{test}} = \frac{1}{n} \sum_{i=1}^{n} L\left(\hat{y}^{[i]}, y^{[i]}\right)$$

ML Terminology (Part 2)

Loss function: Often used synonymously with cost function; sometimes also called error function. In some context the loss for a single data point, whereas the cost function refers to the overall (average or summed) loss over the entire dataset. Sometimes also called empirical risk.

Other Metrics in Future Lectures

- Accuracy (1-Error)
- ROC AUC
- Precision
- Recall
- (Cross) Entropy
- Likelihood
- Squared Error/MSE
- L-norms
- Utility
- Fitness
- ...

But more on other metrics in future lectures.

ML Terminology (Part 3)

Hypothesis: A hypothesis is a certain function that we believe (or hope) is similar to the true function, the target function that we want to model.

Model: In the machine learning field, the terms hypothesis and model are often used interchangeably. In other sciences, they can have different meanings.

Learning algorithm: Again, our goal is to find or approximate the target function, and the learning algorithm is a set of instructions that tries to model the target function using our training dataset. A learning algorithm comes with a hypothesis space, the set of possible hypotheses it explores to model the unknown target function by formulating the final hypothesis.

Classifier: A classifier is a special case of a hypothesis (nowadays, often learned by a machine learning algorithm). A classifier is a hypothesis or discrete-valued function that is used to assign (categorical) class labels to particular data points