

# Lecture 2: Introduction to ML and Classical Models

ML Instruction Team, Fall 2022

CE Department  
Sharif University of Technology

# Machine Learning: An Overview

What is Machine Learning?

# Machine Learning: An Overview

■ Let's review some inspirational quotations ...

▶ *“Machine learning is the hot new thing”*

— John L. Hennessy, President of Stanford (2000–2016)

▶ *“A breakthrough in machine learning would be worth ten Microsofts”*

— Bill Gates, Microsoft Co-Founder

▶ *“Computers are able to see, hear and learn. Welcome to the future.”*

— Dave Waters, Professor at University of Oxford

▶ *“If software ate the world, models will run it”*

— Steven A. Cohen and Matthew W. Granade, The Wallstreet Journal,  
2018

▶ ...

# Machine Learning: An Overview

- The main motivation which we develop (computer) programs is to automate various kinds of (often tedious) processes.
- So far, we have learned to program the computers. the analogy that we are using, is something similar to this:

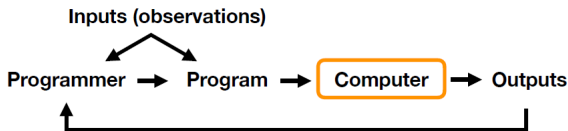


Figure: Classical Programming Paradigm [1].

# Machine Learning: An Overview

- The preceding traditional programming paradigm has several disadvantages:
  - ▶ what if we don't know what program should we write for the given data (inputs) ?
  - ▶ what if the inputs change dynamically over the time? should we write another program?
- In order to resolve such problems, we should replace the need of developing computer programs "manually"
- In other words, we would like to automate the process of creating programs by informing the computer, the inputs and outputs that it needs:



Figure: ML Paradigm [1].

# Categories of Machine Learning

■ The three broad categories of ML are summarized in:

- ▶ **Supervised Learning**
- ▶ **Unsupervised Learning**
- ▶ **Reinforcement Learning**

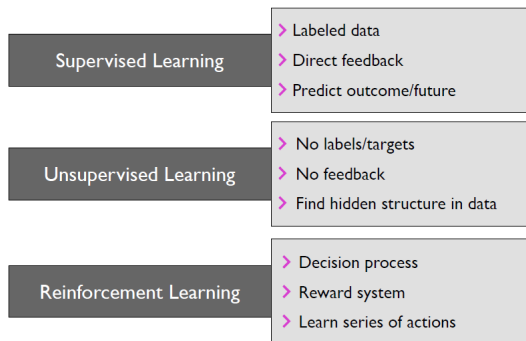


Figure: Categories of ML [2].

# Introduction to Supervised Learning

- Supervised learning is the subcategory of machine learning that focuses on learning from labeled training data, our main goal in supervised learning is summarized in one of these categories:

- **Classification:** predicting the discrete values such as male/female, etc.
- **Regression:** predicting the continuous values such as price, age, etc.

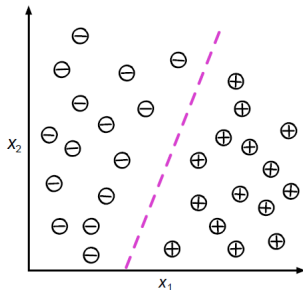


Figure: Illustration of Classification Problem [2].

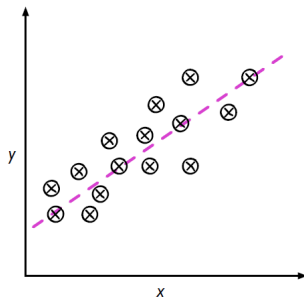


Figure: Illustration of Regression Model [2].

# Supervised Learning

- Given a data set  $\mathcal{D} = \{\langle \mathbf{x}_1, y_1 \rangle, \langle \mathbf{x}_2, y_2 \rangle, \dots, \langle \mathbf{x}_n, y_n \rangle\}$ , there exists an unknown function called  $f$  which:

$$y = f(\mathbf{x})$$

- The supervised learning final goal is to **Approximate** this unknown function. we call our discovery function a *hypothesis* and we define it:

$$\begin{cases} h : \mathbb{R}^m \rightarrow \mathbb{R} \\ h(\mathbf{x}) = y \end{cases}$$



# Unsupervised Learning

- In contrast to supervised learning, unsupervised learning is a branch of machine learning that is concerned with unlabeled data. Common tasks in unsupervised learning are **Clustering** analysis and **Dimensionality Reduction**.

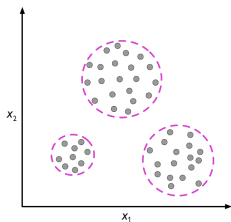


Figure: Illustration of Clustering [2].

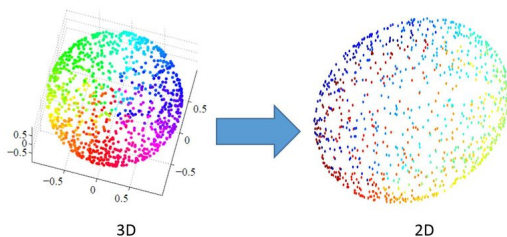


Figure: Illustration of Dimensionality Reduction [3].

# Reinforcement Learning

- Reinforcement is the process of learning from rewards while performing a series of actions.

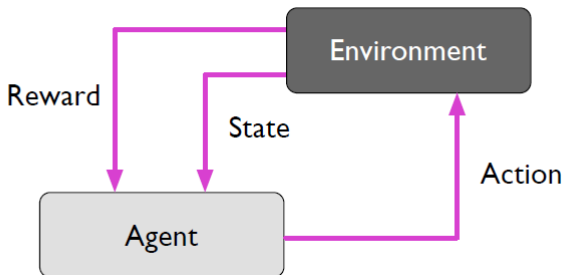
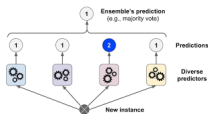


Figure: Illustration of Reinforcement Learning [2].

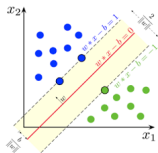
# Classes of Machine Learning Algorithms

- Generalized linear models (e.g., logistic regression)
- Support vector machines (e.g., linear SVM, RBF-kernel SVM)
- Artificial neural networks (e.g., multi-layer perceptrons)
- Tree- or rule-based models (e.g., decision trees)
- Graphical models (e.g., Bayesian networks)
- Ensembles (e.g., Random Forest)
- Instance-based learners (e.g., K-nearest neighbors)

# Classes of Machine Learning Algorithms



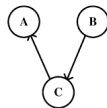
(a) Ensemble Learning [4].



(b) Support Vector Machine [7].

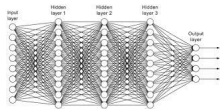


(c) Decision Tree [6].

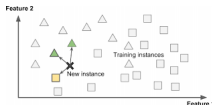


$$P(A,B,C) = P(A|C) \cdot P(B|C) \cdot P(C)$$

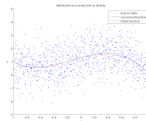
(d) Graphical Models [8].



(e) Neural Networks [9].



(f) K-Nearest Neighbors [12].



(g) Generalized Linear Models [5].

# Algorithm Categorization Schemes

- Eager vs Lazy
- Single-Task vs Multi-Task
- Generative vs Discriminant
- Instance-based vs Model-Based
- Parametric vs Non-Parametric
- Batch vs Online

# 5 Steps To Solve A Machine Learning Problem

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

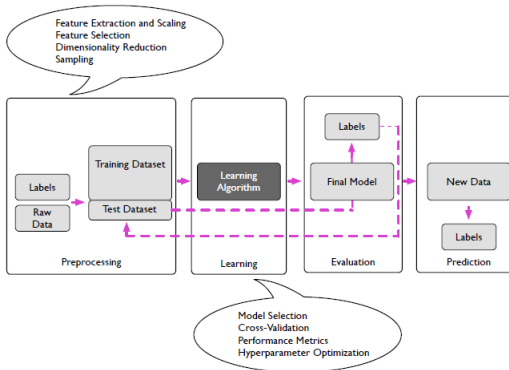


Figure: Learning Process [2].

# 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

- Combinatorial search, greedy search (e.g., decision trees over, not within nodes);
- Unconstrained convex optimization (e.g., logistic regression);
- Constrained convex optimization (e.g., SVM);
- Nonconvex optimization, here: using backpropagation, chain rule, reverse autodi. (e.g., neural networks).
- Constrained nonconvex optimization (semi-adversarial networks, not covered in this course)

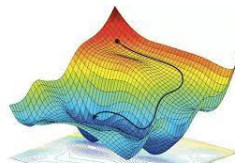


Figure: Gradient Descent [10].

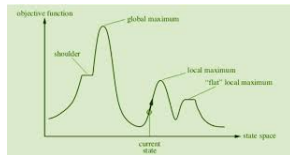


Figure: Hill Climbing [11].



# Evaluation

- There are several different evaluation metric to assess the performance of a model, some of them are:
  - ▶ Accuracy (1-Error)
  - ▶ ROC AUC
  - ▶ Precision
  - ▶ Recall
  - ▶ (Cross) Entropy
  - ▶ Likelihood
  - ▶ Mean Squared Error (MSE)
  - ▶ Mean Absolute Error (MAE)
  - ▶ L-norms
  - ▶ ...

# Glossary

- **Training example:** A row in the table representing the dataset.
- **Training:** Model fitting, for parametric models similar to parameter estimation.
- **Feature,  $x$ :** A column in the table representing the dataset.
- **Predicted output,  $\hat{y}$ :** Use this to distinguish from targets; here, means output from the model.
- **Loss function:** Often used synonymously with cost function.
- **Hypothesis::** A hypothesis is a certain function that we believe (or hope) is similar to the true function.
- **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.
- **Hyperparameters:** Hyperparameters are the *tuning parameters* of a machine learning algorithm.
- **Model:** In the machine learning field, the terms *hypothesis* and *model* are often used interchangeably. In other sciences, they can have different meanings: A hypothesis could be the "educated guess" by the scientist, and the model would be the manifestation of this guess to test this hypothesis.
- **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.

# References

- [1]. Raschka, Sebastian. “What Are Data Science and Machine Learning?” Dr. Sebastian Raschka, 3 Sept. 2022, [sebastianraschka.com/faq/docs/datascience-ml.html](https://sebastianraschka.com/faq/docs/datascience-ml.html).
- [2]. Raschka, Sebastian, and Vahid Mirjalili. Python Machine Learning: Machine Learning and Deep Learning With Python, Scikit-learn, and TensorFlow 2, 3rd Edition. 3rd ed., Packt Publishing, 2019.
- [3]. Peluffo, Diego. Dimensionality Reduction Effect Over an Artificial (3-dimensional) Spherical Shell Manifold. Resultant Embedded (2-dimensional) Data Is an Attempt to Unfolding the Original Data. Feb. 2017, [www.researchgate.net/publication/313787026-Interactive-Data-Visualization-Using-Dimensionality-Reduction-and-Similarity-Based-Representations](https://www.researchgate.net/publication/313787026-Interactive-Data-Visualization-Using-Dimensionality-Reduction-and-Similarity-Based-Representations).
- [4]. Kumar, Ajitesh. “5 Common Ensemble Methods in Machine Learning.” Data Analytics, 16 Aug. 2022, [vitalflux.com/5-common-ensemble-methods-in-machine-learning](https://vitalflux.com/5-common-ensemble-methods-in-machine-learning).
- [5]. <http://strijov.com/sources/demo-GLM.php>
- [6]. [www.researchgate.net/figure/Schematic-of-a-Decision-Tree-The-figure-shows-an-example-of-a-decision-tree-with-3-fig1-348456545](https://www.researchgate.net/figure/Schematic-of-a-Decision-Tree-The-figure-shows-an-example-of-a-decision-tree-with-3-fig1-348456545). Accessed 8 Sept. 2022.

# References

- [7]. Wikipedia contributors. “Support-vector Machine.” Wikipedia, 1 Sept. 2022, [en.wikipedia.org/wiki/Support-vector-machine](https://en.wikipedia.org/wiki/Support-vector-machine).
- [8]. [www.researchgate.net/figure/Simple-directed-graphical-model-with-three-variables-To-illustrate-how-graphical-models-fig6-262407302](https://www.researchgate.net/figure/Simple-directed-graphical-model-with-three-variables-To-illustrate-how-graphical-models-fig6-262407302). Accessed 8 Sept. 2022.
- [9]. Ashtari, Hossein. “What Is a Neural Network? Definition, Working, Types, and Applications in 2022.” Spiceworks, 3 Aug. 2022, [www.spiceworks.com/tech/artificial-intelligence/articles/what-is-a-neural-network](https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-a-neural-network).
- [10]. Balaouras, Georgios. “Optimization Algorithms.” Georgios Balaouras, 21 Apr. 2022, [mpalaourg.me/project/optimization-algorithms](https://mpalaourg.me/project/optimization-algorithms).
- [11]. GeeksforGeeks. “Introduction to Hill Climbing | Artificial Intelligence.” GeeksforGeeks, 23 Aug. 2022, [www.geeksforgeeks.org/introduction-hill-climbing-artificial-intelligence](https://www.geeksforgeeks.org/introduction-hill-climbing-artificial-intelligence).
- [12]. Agrawal, Sanidhya. “What Is Instance-Based Learning? - Sanidhya Agrawal.” Medium, 14 Dec. 2021, [medium.com/@sanidhyaagrawal08/what-is-instance-based-learning-a9b06079e836](https://medium.com/@sanidhyaagrawal08/what-is-instance-based-learning-a9b06079e836).

**Thank You!**

**Any Question?**