CSE 411: Machine Learning

Introduction

Dr Muhammad Abul Hasan

Department of Computer Science and Engineering
Green University of Bangladesh
muhammad.hasan@cse.green.edu.bd

Fall 2023

Outline

“ Wisdom is not a product of schooling but of the lifelong attempt to acquire it.

– Albert Einstein *”*

Data and Learning

Big Data

- Widespread use of personal computers and wireless communication leads to "Big Data"
- We are both *producers* and *consumers* of data
- Data is not random, it has structure, e.g., customer behavior
- We need "big theory" to extract that structure from data for:
  1. Understanding the process
  2. Making predictions for the future

Why "Learn"?

- Machine learning is programming computers to optimize a performance criterion using example data or past experience.
- There is no need to "learn" to calculate payroll
- Learning is used when:
  - Human expertise does not exist (navigating on Mars),
  - Humans are unable to explain their expertise (speech recognition)
  - Solution changes in time (routing on a computer network)
  - Solution needs to be adapted to particular cases (user biometrics)

What We Talk About When We Talk About "Learning"

- Learning general models from a data of particular examples
- Data is cheap and abundant (data warehouses, data marts); knowledge is expensive and scarce.
- Example in retail: Customer transactions to consumer behavior:
  - People who buy "Toothbrush" also buy "Toothpaste" (e-commerce sites)
- Build a model that is a good and useful approximation to the data.

Data Mining

- **Retail**: Market basket analysis, Customer relationship management (CRM)
- **Finance**: Credit scoring, fraud detection
- **Manufacturing**: Control, robotics, troubleshooting
- **Medicine**: Medical diagnosis
- **Telecommunications**: Spam filters, intrusion detection
- **Bioinformatics**: Motifs, alignment
- **Web mining**: Search engines
- ...

Machine Learning

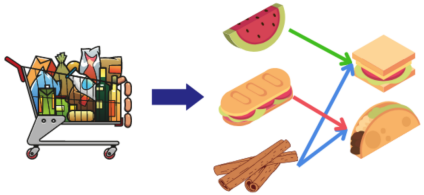
What is Machine Learning?

- Optimize a performance criterion using example data or past experience.
- Role of Statistics: Inference from a sample
- Role of Computer science: Efficient algorithms to:
  - Solve the optimization problem
  - Representing and evaluating the model for inference

Applications

- Association
- Supervised Learning
  - Classification
  - Regression
- Unsupervised Learning
- Reinforcement Learning

Learning Associations



"93% of people who purchased item A also purchased item B"

- Basket analysis: $P(Y|X)$ probability that somebody who buys X also buys Y where X and Y are products/services.
  - Example: $P(chips|beer) = 0.7$

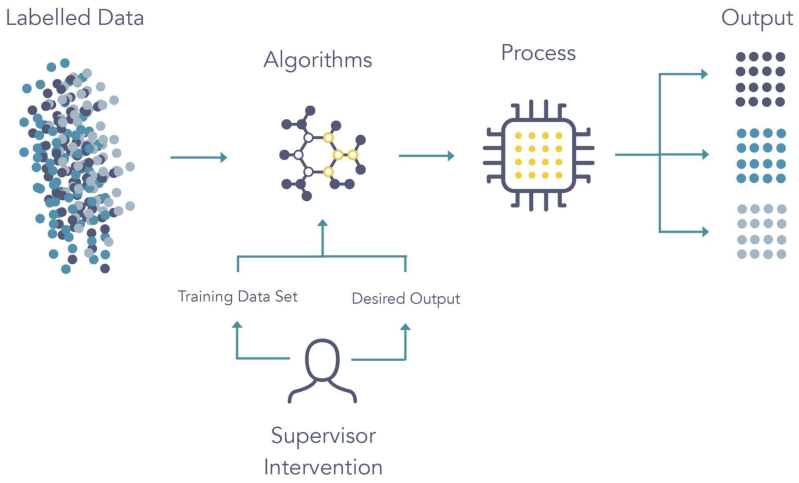
Learning Association: Question

Question

In basket analysis, we want to find the dependence between two items X and Y. Given a database of customer transactions, how can we find these dependencies? How would we generalize this to more than two items?

Supervised Learning

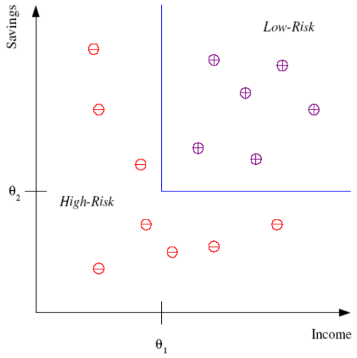
Supervised Learning



Supervised Learning: Classification

Example: Credit scoring

- Differentiating between low-risk and high-risk customers from their income and savings



Discriminant: IF income $> \theta_1$ AND savings $> \theta_2$ THEN low-risk ELSE high-risk

Classification: Applications

- *Also known as*, Pattern recognition
- **Face recognition**: Pose, lighting, occlusion (glasses, beard), make-up, hairstyle
- **Character recognition**: Different handwriting styles.
- **Speech recognition**: Temporal dependency.
- **Medical diagnosis**: From symptoms to illnesses
- **Biometrics**: Recognition/authentication using physical and/or behavioral characteristics: Face, iris, signature, etc
- **Outlier/novelty detection**:

Classification: Question

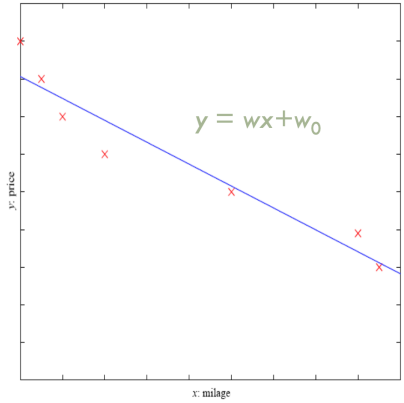
Question

In a daily newspaper, find five sample news reports for each category of politics, sports, and the arts. Go over these reports and find words that are used frequently for each category, which may help you discriminate between different categories. For example, a news report on politics is likely to include words such as "government," "recession," "congress," and so forth, whereas a news report on the arts may include "album," "canvas," or "theater." There are also words such as "goal" that are ambiguous.

Supervised Learning: Regression

Example: Price of a used car

- x : car attributes
- y : price
- $y = g(x|q)$
  - $g()$ model
  - q parameters



Regression: Question

Question

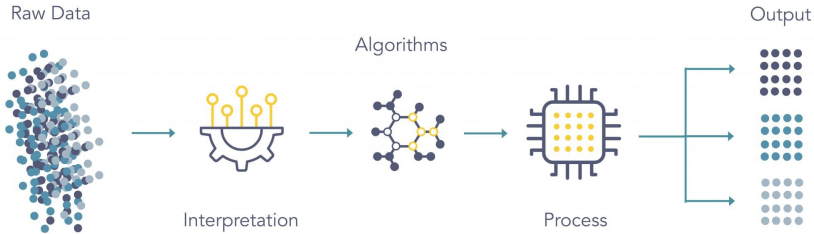
In estimating the price of a used car, it makes more sense to estimate the percent depreciation over the original price than to estimate the absolute price. Why?

Supervised Learning: Uses

- **Prediction of future cases**: Use the rule to predict the output for future inputs
- **Knowledge extraction**: The rule is easy to understand
- **Compression**: The rule is simpler than the data it explains
- **Outlier detection**: Exceptions that are not covered by the rule, e.g., fraud

Unsupervised Learning

Unsupervised Learning

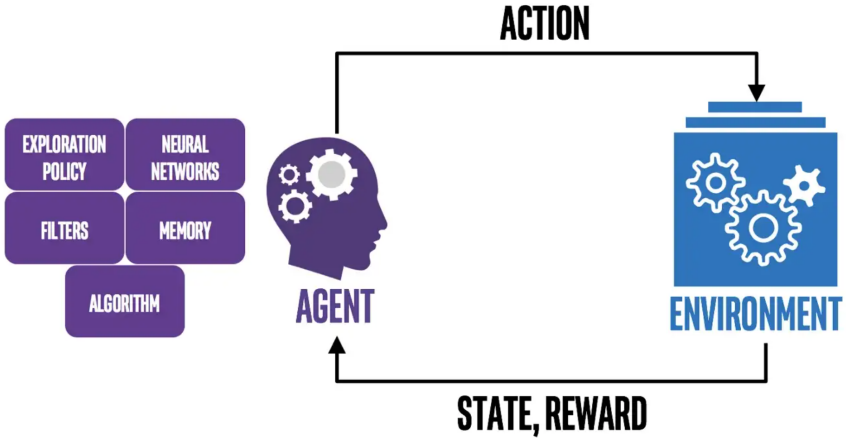


Unsupervised Learning

- Learning "what normally happens"
- **Clustering**: Grouping similar instances
- Example applications
  - **Genetics**: for example, clustering DNA patterns to analyze evolutionary biology.
  - **Recommender systems**: which involve grouping together users with similar viewing patterns in order to recommend similar content.
  - **Anomaly detection**: including fraud detection or detecting defective mechanical parts.

Reinforcement Learning

Reinforcement Learning



Reinforcement Learning

- Learning a *policy*: A sequence of outputs
- No supervised output but delayed reward
- Applications:
  - Credit assignment problem
  - Game playing
  - Robot in a maze
  - Multiple agents, partial observability, ...

Reinforcement Learning: Question

Question

How would you approach a computer programming competition?

Thank You!