Machine Learning

Introduction to Machine Learning

Slides adapted from material created by E. Alpaydin
Prof. Mordohai, Prof. Greenstadt, Pattern Classification (2nd Ed.),
Pattern Recognition and Machine Learning

Objectives

- Understand common machine learning problems
- Understand basic ML terminology

What is Machine Learning?

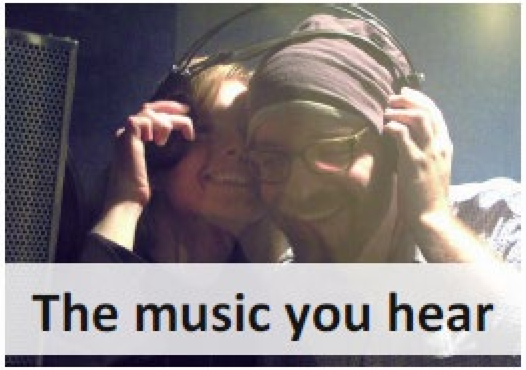
- Definition: “The study of computer algorithms that improve automatically through experience”
- Formally:
  - Improve at task T
  - With respect to performance measure P
  - Based on experience E
- Example: Recognize a Person
  - T: recognize a person
  - P: number of time we recognized a person correctly
  - E: a database of labeled faces

ML vs AI

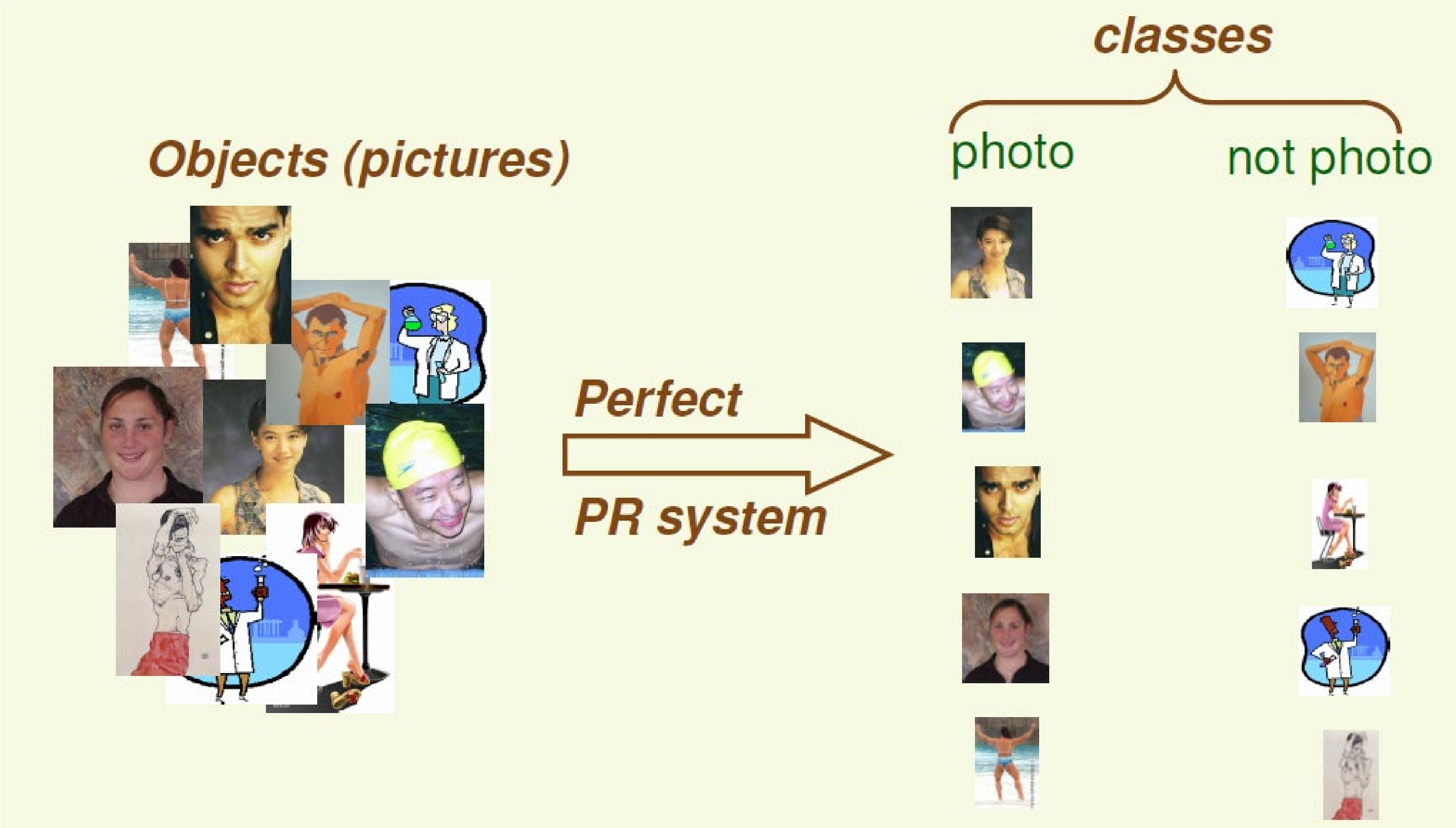
- How is this different than AI?
  - ML can be thought of as a sub-topic within AI
- AI deals with any “intelligent” task performed by a non-human agent
  - Often “path finding algorithms”
- ML specifically deals with making decisions based on acquired data
  - Both past and current

Why do we care?

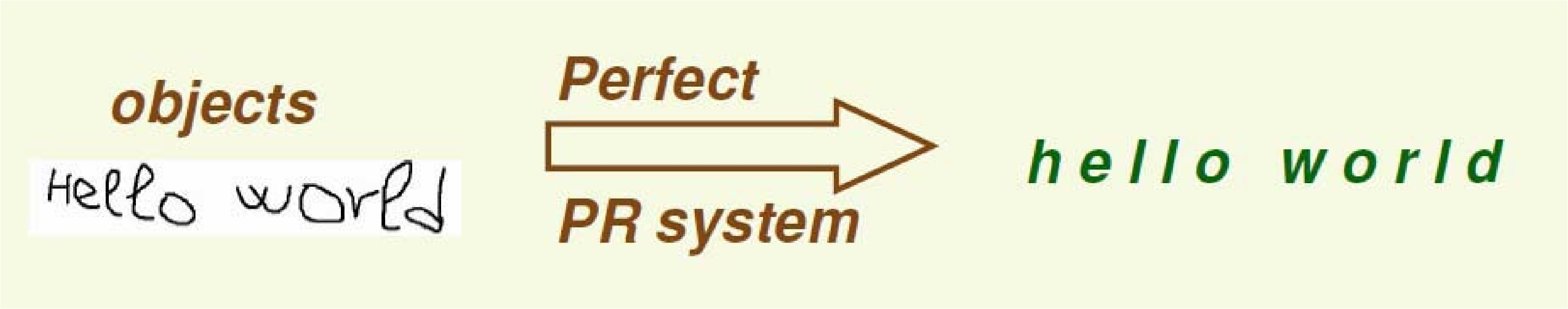
- It’s everywhere!!!



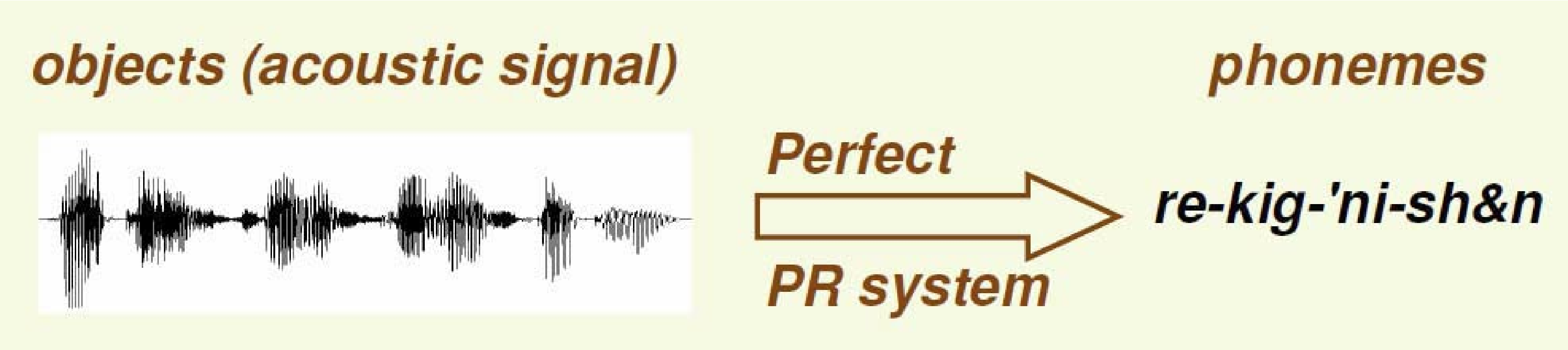
Example: Photograph or Not?



Example: Character Recognition



Example: Speech Understanding

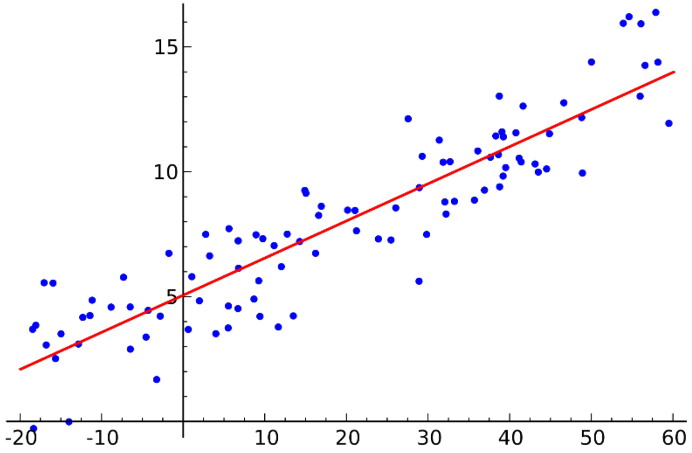


How Do We Do It?

- We need to think about
  1. What needs to be learned?
     - What’s our task/goal?
  2. What feedback can we get and in what form?
     - Supervised learning (correct answers for each example)
     - Unsupervised learning (correct answers not given. **Not covered in this class**)
  3. What representation should we use (features)?

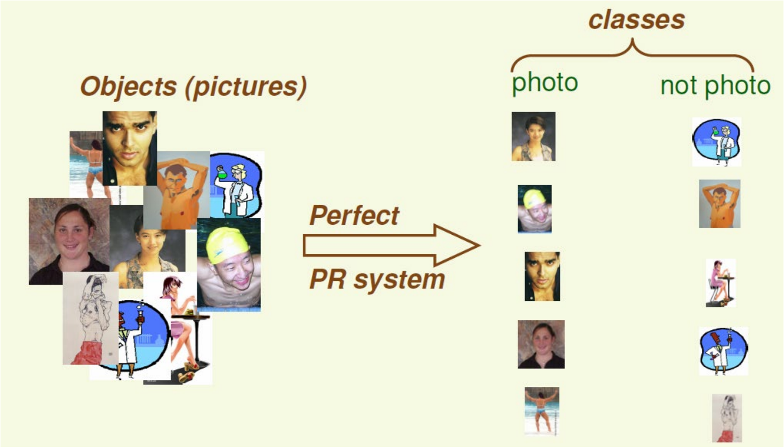
Problems in ML

- There are two types of machine learning problems that we'll tackle in the course
- Regression
  - Given some data, can we predict an outcome value?
  - Example: We have a car's brand, year, mpgs and want to figure out its worth
  - This is an example of *supervised* learning
    - To build our prediction system, we have data with labels.



Problems in ML

- Classification
  - Given data, can we predict which category something belongs to.
  - Typically involves *learning* some rules.
  - This is also an example of *supervised* learning
    - To build our prediction system, we have data with labels.



Course Objectives

- Foundations of Machine Learning
  - Regression
  - Classification
- Applications of Machine Learning algorithms
- Implementation and use of Machine Learning algorithms
- Note: This course should probably be called something like “Foundations of Traditional Machine Learning”

CCI AI Course Offerings

Computer Science

- CS 380/510 – Artificial Intelligence
- **CS 383/613 – Machine Learning**
- CS 387/611 – Game AI Development
- CS 481/610 – Advanced Artificial Intelligence
- CS 482/589 – Robust Machine Learning
- CS 486/770 – Topics in Artificial Intelligence
- CS 614 – Applications of Machine Learning
- CS 615 – Deep Learning
- CS 616 – Robust Deep Learning
- CS 617 – Reinforcement Learning

Information Science/Data Science

- DSCI 471 – Applied Deep Learning
- DSCI 631 - Applied Machine Learning for Data Science
- DSCI 691 – Natural Language Processing with Deep Learning
- INFO 629 – Applied Artificial Intelligence
- INFO 692 – Explainable Artificial Intelligence
- INFO 693 – Human-Artificial Intelligence Interaction

Administrative Stuff…

Contacts

Instructor

- Professor Matt Burlick:
  - Email: mjburlick@drexel.edu
  - Office:
    - 3675 Market St., Room 925
  - Office Hours:
    - Mondays 01:00pm – 03:00pm (Hybrid)
    - Thursdays 01:00pm – 03:00pm (Hybrid)

Teacher Assistants

- None ☹

Contacts

- Questions are to be asked during class or office hours.
  - Time is allocated for questions at both the beginning and end of class time.
- Outside of logistical questions, questions will not be answered over email or discord.
- Discord will be used for student-led discussions (although monitored by faculty) and for announcements.

Pre-Requisites

- CS 260 [Min Grade: C] – Data Structures and Algorithms
- MATH 201 [Min Grade: C] or ENGR 231 [Min Grade: D] – Linear Algebra
- MATH 221 [Min Grade: C] or MATH 222 [Min Grade: C] – Discrete Math
- MATH 311 [Min Grade: C] or MATH 410 [Min Grade: C] or ECE 361 [Min Grade: D] – Prob and Stats
- The idea is that you should be a proficient programmer such that you can pick up a new language "on the fly" and use it as a tool.
- You should also be comfortable with linear algebra, probability, statistics, and calculus.

Course Resources

- Official Textbook:
  - None
- Recommended Textbooks:
  - Basic: An Introduction to Statistical Learning (free PDF), Gareth James, et. al.
  - Medium: Introduction to Machine Learning (Alpaydin)
  - Advanced: Machine Learning (Murphy)
- Blackboard for lecture material, labs and assignments
- Discord Channel

CS383-Sp25

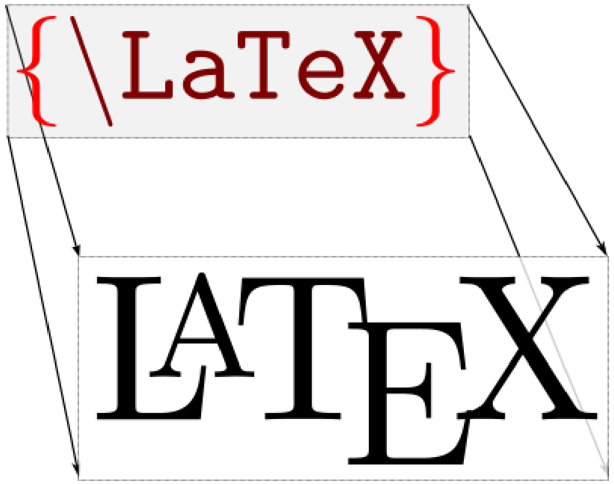
  - Use this as your first place to pose questions
    - Hopefully not just I can help
  - But don’t post code.

Course Software

- Programming Environment: Python 3.x w/
  - NumPy
  - MatplotLib
  - Opencv-python
  - Pandas
  - Pillow

Course Software

- Typesetting Environment:
  - LaTeX, MS Word w/ Equation Editor (or similar)
  - If you opt to do LaTeX, unless you already have a LaTeX environment set up, I recommend using an online LaTeX typesetter, ala www.overleaf.com



Grade Breakdown

- Labs 30%
- Homework Assignments 20%
- Exams 50%

Points	Grade	Points	Grade	Points	Grade
[97-100]	A+	[83-98)	B	[70-73)	C-
[93-97)	A	[80-83)	B-	[67-70)	D+
[90-93)	A-	[77-80)	C+	[60-67)	D
[87-90)	B+	[73-77)	C	[0-60)	F

Labs

- Most weeks there will be a **lab** conducted during the second lecture.
- Lab attendance is **required**
  - But you can miss up to two without penalty.
  - Subsequent will result in zeros.
- You may work with another student if you like.
- Your lowest lab grade will be dropped.
- No late labs will be accepted.

Assignments

- Most weeks there will be a **homework assignment.**
  - These focus on the theory and math.
  - These are to be typeset via LaTeX or Microsoft Equation Editor (or similar) and converted to PDF.
- Your lowest assignment grade will be dropped.
- No late assignments will be accepted.

Exams

- There will be a midterm and final exam.
- These will be similar in style to the questions in the assignments.

Use of AI Tools

- Artificial intelligence tools such as large language models (e.g., ChatGPT) are permitted to be used as a reference in studying and understanding labs and assignments.
- HOWEVER, you may not use AI tools to generate solutions (code and/or computations) for your work.

Additional Course Policies

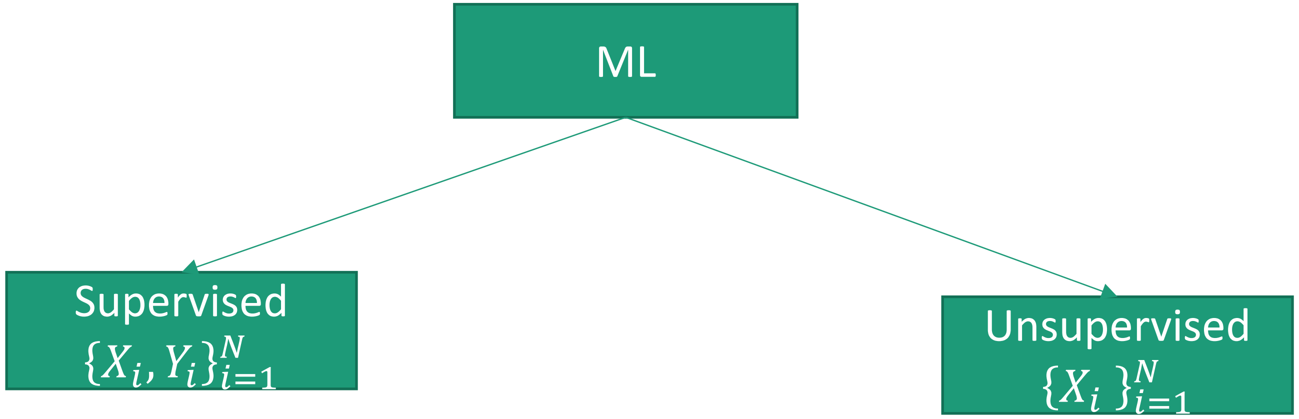
- Assignments and Exams are to be done individually unless otherwise noted
- While you are encouraged to use a versioning system like github or bitbucket, please make your work for this course **private.**
- Any dispute about an assignment grade must be formally made (email) and resolved within 5 days of receiving your grade. After this period your grade cannot be adjusted.

Notation/Mathematics/Matlab

- I have placed on Blackboard a number of resources:
  - Course Notation – There will be a lot of symbols used in this course. This document tries to give you an overview of them.
  - Similarity and Distance Functions – Often we will need to compute the distance and/or similarity between observations. This document includes several commonly used ones.
  - Python Functions – Here's a list of most of the Python functions I used in developing this course.
  - Math Review – A quick review of the most critical math needed for this course. Including..
    - Calculus
    - Linear Algebra
    - Probability and Statistics
  - Math Reference Sheet – This is a quick reference sheet that we'll use often when doing derivations and whatnot.

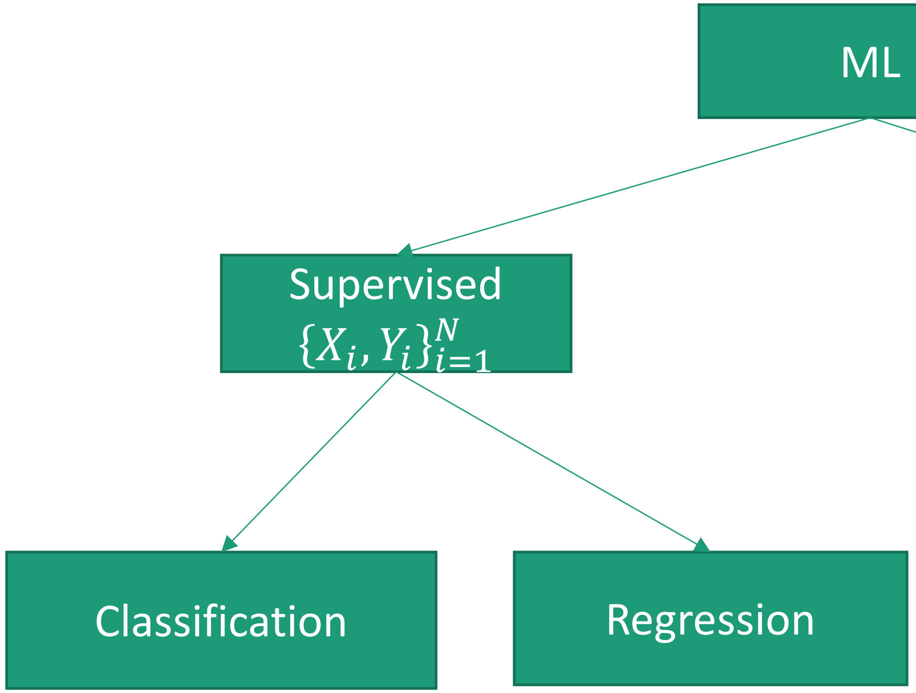
Fundamental ML Concepts

ML Overview

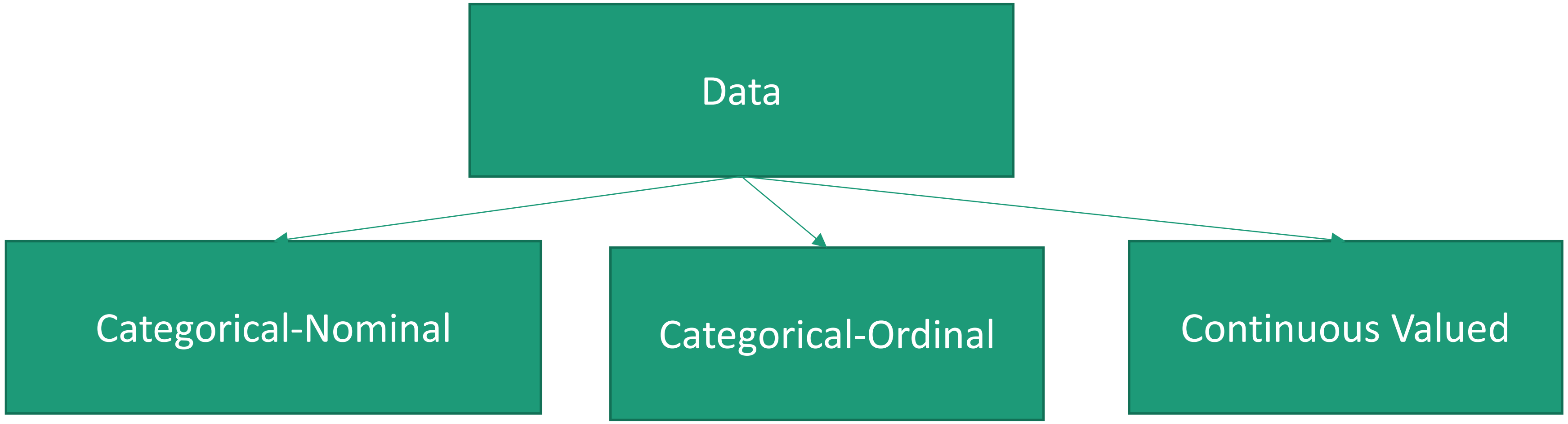


- We can basically break machine learning tasks into two categories
  1. Supervised Learning
  2. Unsupervised Learning
- *Supervised learning*
  - Data X_i and correct answer (label) Y_i given for each example $i \in \{1, \dots, N\}$
- *Unsupervised learning*
  - Only data given for each example, X_i
- Again, we'll be just focusing on *supervised learning.*

Types of Problems



Types of Data



- Each piece of information pertaining to an observation can fall into one of three categories:
  - *Continuous Valued*
    - Examples: Blood Pressure, Height
  - *Categorical-Nominal (unordered)*
    - Examples: Car Model, School
  - *Categorical-Ordinal (can be ordered)*
    - Examples: Colors, small < medium < large

No Free Lunch Theorem

- Unfortunately, there’s no single machine learning algorithm to rule them all ☹
- Hopefully, the nature of the problem and data will guide us towards some subset of the options.
- We then try them out and select the best.

ML Algorithms

- Here's a list of algorithms we'll look at in the class and what types they are

1. Principal Component Analysis (PCA)
2. Linear Regression
3. Classification
    a. Binary
        a. Logistic Regression
        b. Support Vector Machines (SVMs)
    b. Multi-Class
        a. Decision Trees (DTs)
        b. Nearest Neighbors (KNN)
        c. Statistical Classification
        d. Markov Models