

# A VERY BRIEF INTRODUCTION TO MACHINE LEARNING

*Abdus Salam Azad*

# What Is Machine Learning

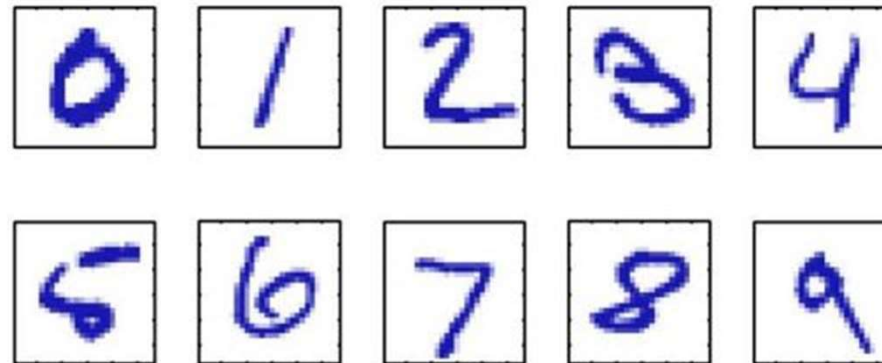
- Tom Mitchell - Improving with experience at some task
- Wikipedia - **Machine learning** is a [scientific discipline](#) that explores the construction and study of [algorithms](#) that can *learn from data*
- **Algorithms that can improve their performance using data**

# What Is Machine Learning

- Creating programs that can automatically learn rules from data

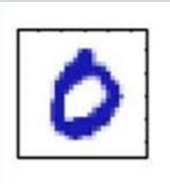
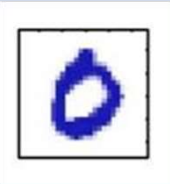

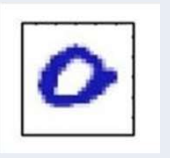


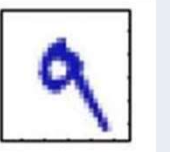
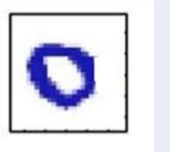
# Example – Character Recognition(Classification)

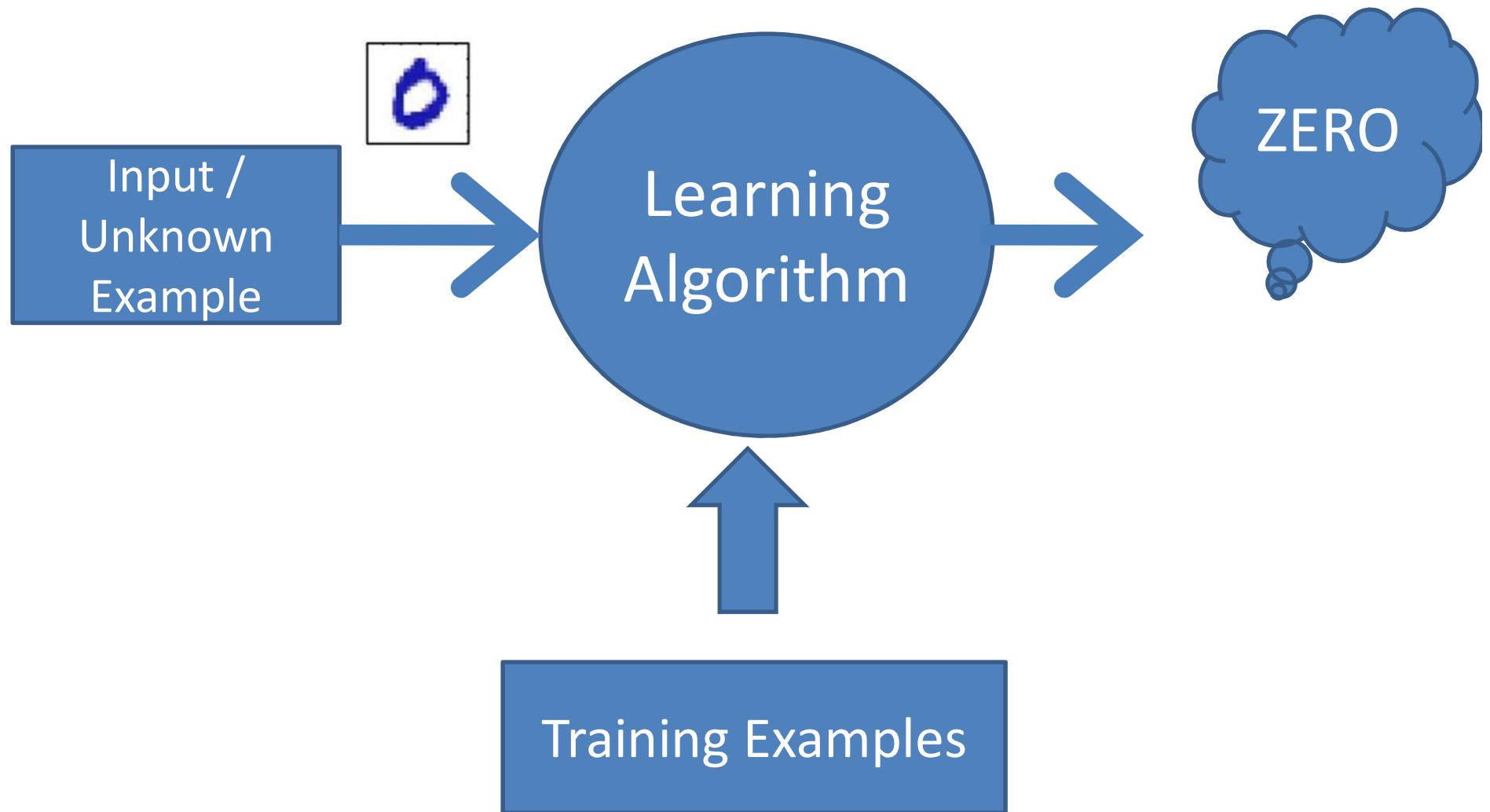
- Lets say we have 1000 *worked-out* examples
- We can then **train** an algorithm with these examples, which can then recognize **unseen** examples



Images are 28 x 28 pixels

# Training Examples

Image	Classification	Image	Classification
	Zero		Zero
	Three		Zero
	Five		Zero
	Nine		Zero



# Enjoy Sports (Classification)

$\langle ?, \text{Cold}, \text{High}, ?, ?, ? \rangle$

Example	<i>Sky</i>	<i>AirTemp</i>	<i>Humidity</i>	<i>Wind</i>	<i>Water</i>	<i>Forecast</i>	<i>EnjoySport</i>
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

**TABLE 2.1**

Positive and negative training examples for the target concept *EnjoySport*.

# Function

## Approximation/Regression

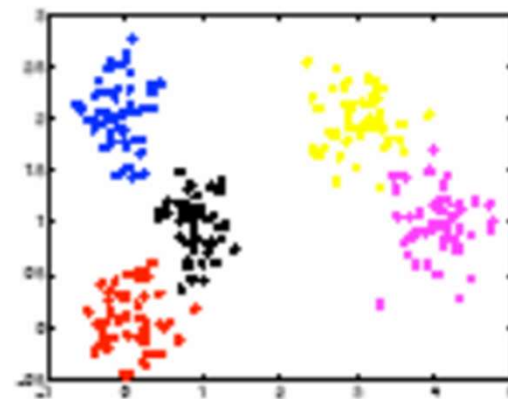
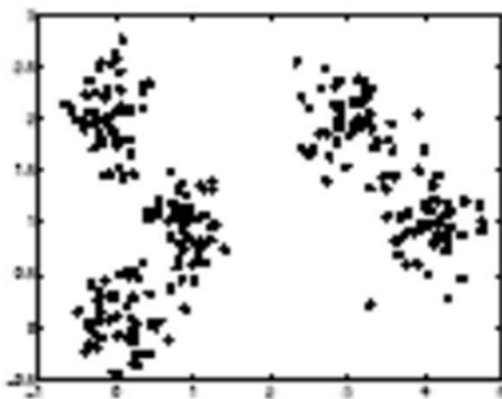
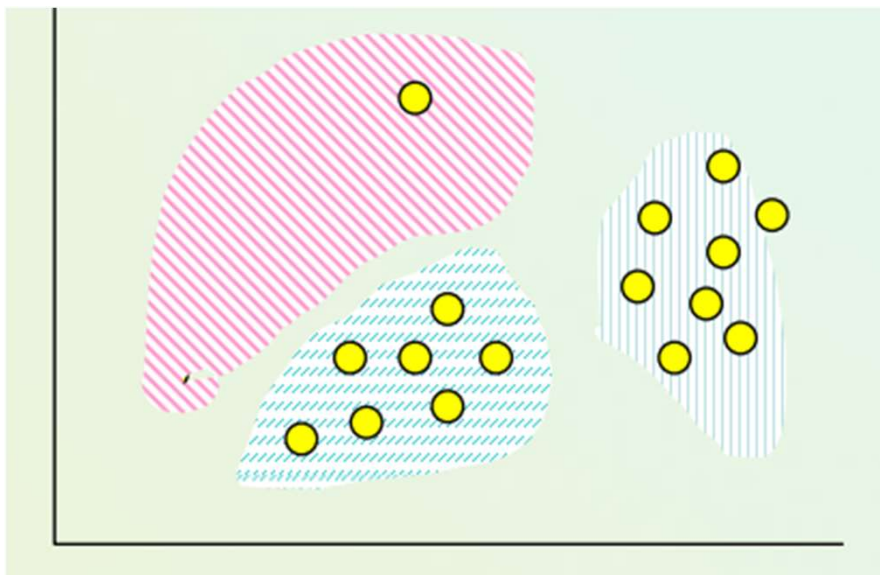
- We have information about Tamim Iqbal's last 100 One Day innings

Weather	Venue	Pitch	Day/Night	...	Run
Cold	Mirpur	Flat	Yes	...	15
Mild	Auckland	Green	Yes	...	54
Sunny	Lords	Dusty	No	...	83
...	...	...	...	...	...
Cold	Mirpur	Flat	No	...	97

- What will be Tamim Iqbal's score if?
  - Weather – Mild, Pitch – Flat, Venue-Dhaka D/N- No



# Clustering



# Clustering

- Find group of *similar* points
- Help marketers discover distinct groups in their customer bases, and then use this knowledge to develop targeted marketing programs
- Each customer may be represented by their salary, amount of products bought per month, living place, education, qualification, number of credit cards used, etc.

# Learning

- Learning can be classified as follows
  - Supervised Learning
  - Unsupervised Learning
  - Semi-Supervised Learning

# Supervised Learning

- The agent observes some example input output pairs and learns a function that maps from input to output
- Two basic types
  - Classification
  - Regression

# Supervised Learning

- Labeled Examples
- Given input-output pairs as example
- Learns a function/hypothesis that maps input to output
  - Can be thought as there is a teacher/oracle who gives you “labeled examples”

**Find a problem that falls under  
supervised learning**



# Supervised Learning Scenarios

- Face Recognition
- Character Recognition
- Document Classification
- Sentiment Analysis
- Disease Detection

# Unsupervised Learning

- the agent learns patterns in the input even though no explicit feedback is supplied
- The most common unsupervised learning task is **clustering**



# Unsupervised Learning Scenarios

# Reinforcement Learning

- the agent learns from a series of reinforcements
  - rewards or punishments

**What is the difference between  
reinforcement learning and  
supervised learning ????**

# Three Main Types of Learning

- The three main types arise based on the type of feedback that the agent gets from the environment

# The importance of feature vector in case of supervised learning

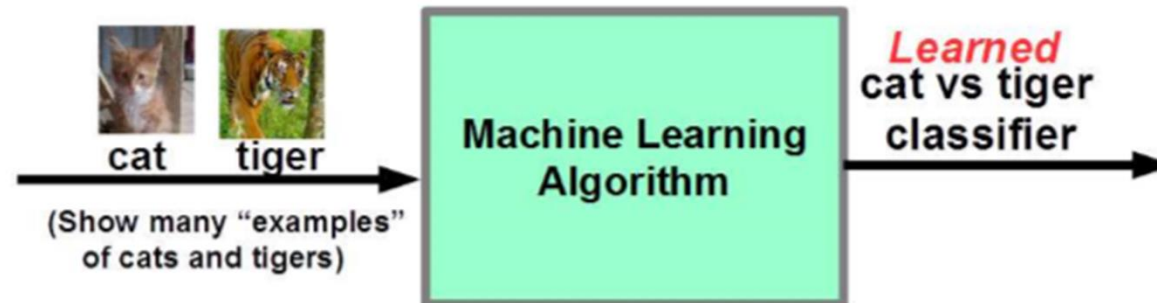
- The features should actually have effect on the classification

# Traditional algorithms vs Machine Learning algorithms

- Traditional: Write programs using hard-coded (fixed) rules



- Machine Learning (ML): **Learn rules** by looking at some **training data**



# Some real-world applications

- Information retrieval (text, visual, and multimedia searches)
- Machine Translation
- Question Answering
- Social networks
- Recommender systems (Amazon, Netflix, etc.)

# Some real-world applications

- Speech/handwriting/object recognition
  - Ad placement on websites
  - Credit-card fraud detection
  - Weather prediction
  - Autonomous vehicles (self-driving cars)
  - Healthcare and life-sciences
- ... and many more applications in sciences and engineering



# Data Sets

- <http://archive.ics.uci.edu/ml/>
  - Old : <http://kdd.ics.uci.edu/>

# Resource

- Google
- Machine Learning by Tom Mitchell

**Any Questions**



