Machine Learning

AIML CZG565

M1 : Introduction

BITS Pilani

Pilani Campus

Dr. Pankaj Agarwal, Professor & Guest faculty

Profile : Dr. Pankaj Agarwal

innovate achieve lead

Total Experience: More than 23 years of Experience in Academics & Industry.
Current Status: Guest Faculty at BITS Pilani **and** Professor & Dean (Engineering)
Areas of interest & research: Data Science & Machine Learning algorithms, Algorithms Design & Analysis, Data Analytics, Soft computing Techniques, NLP, Computer Vision & Deep Learning with Python Programming, Tableau, KNIME etc
Research Publications: more than 60 research publications in international journals/conferences.
Patents: 10 research patents including Australian, German & Indian patents.
Certifications: IBM certified data Science Professional & Microsoft certified Programmer.
Consultant & Trainer: BITS Pilani, Ureka International, Princeton Hive, Great Learning.
Books: Design & Analysis of Algorithms, .NET,SPM, RDBMS, C programming, Computer Organization, Cyber Security
Research Guidance: more than 100 M.Tech Students (BITS Pilani), 5 Ph.D students completed

Machine Learning

innovate achieve lead

Disclaimer and Acknowledgement

- The content for these slides has been obtained from books and various other source on the Internet
- I hereby acknowledge all the contributors for their material and inputs.
- I have provided source information wherever necessary
- I have added and modified the content flow to suit the requirements of the course and for ease of class presentation
- Students are requested to refer to the textbook and detailed content of this presentation deck over canvas

Course Introduction

innovate achieve lead

- **Objective of course**
  - Introduction to the basic concepts and techniques of Machine Learning
  - Gain experience in basics of doing independent study and research in the field of Machine Learning
  - Develop skills of using recent machine learning software tools to evaluate learning algorithms and model selection for solving practical problems
- **Focus of this course**
  - Strong Mathematical Foundations of ML algorithms
  - Structured Data Analytics
  - IDD (Independent & Identically Distributed Data)
- **Topics not expected of this course**
  - Unstructured Data Analytics
  - Time Series/Sequence Data Analytics
  - Deep Learning

Course Plan

M1	Introduction
M2	Machine learning Workflow
M3	Linear Models for Regression
M4	Linear Models for Classification
M5	Decision Tree
M6	Instance Based Learning
M7	Support Vector Machine
M8	Bayesian Learning
M9	Ensemble Learning
M10	Unsupervised Learning
M11	Machine Learning Model Evaluation/Comparison

Text books and Reference book(s)

innovate achieve lead

T1 Tom M. Mitchell: Machine Learning, The McGraw-Hill Companies

R1 Christopher M. Bishop: Pattern Recognition & Machine Learning, Springer

P. Tan, et al. Introduction to Data Mining, Pearson

R2 C.J.C. BURGES: A Tutorial on Support Vector Machines for Pattern Recognition,

R3 Kluwer Academic Publishers, Boston.

Evaluation scheme

- **Quiz (10% - Best 2 of 3 quizzes)**
- **Assignment (20% - 1 Progressive Group Assignment)**
- **Mid-semester exam (30%)**
- **Comprehensive exam (40%)**

Pre-requisites

- Linear algebra: vector/matrix manipulations, properties
- Calculus: partial derivatives
- Probability: common distributions; Bayes Rule
- Statistics: mean/median/mode; maximum likelihood

Lab Plan

Lab No.	Lab Objective
1	End to End Machine Learning
2	Linear Regression and Gradient Descent Algorithm
3	Logistic Regression Classifier
4	Decision Tree
5	Naïve Bayes Classifier
6	Random Forest

- **Labs not graded**
- **Most of the Lab recordings available at CSIS virtual labs**
- **Webinars will be conducted for lab sessions**
- **Labs will be conducted in Python**

Agenda

- What is Machine Learning?
- Why Machine Learning is important?
- Types of Machine Learning
- Application Areas
- Issues in Machine Learning
- Demo Case study

What is Machine Learning?

- The term Machine Learning was first coined by **Arthur Samuel** in the year 1959, an IBM employee and pioneer in the field of computer gaming and artificial intelligence
- Humans can learn from their experiences with their learning capability,
- But can a machine also learn from experiences or past data like a human does?
- It is the branch of AI mainly concerned with the development of algorithms which allow a computer to learn from the **data** and **past experiences** on their own.
- Machine Learning is a system that can learn from **example** through **self-improvement** and without being explicitly coded by programmer.
- Machine learning uses various algorithms for building **mathematical models** and making **predictions** using historical data or information

ML – What, When, Where ?

What is Machine Learning?

Definition by Tom Mitchell (1998):

"A computer program is said to learn from **experience E** with respect to some class of **tasks T** and **performance measure P**, if its performance at tasks in T, as measured by P, **improves with experience E**."

Example: playing checkers.

E = the experience of playing many games of checkers

T = the task of playing checkers.

P = the probability that the program will win the next game.

What is Machine Learning?

- To have a learning problem, we must identify
  - The class of tasks
  - The measure of performance to be improved
  - Source of experience

Modern-day machine learning has two objectives, one is to **classify data** based on models which have been developed, the other purpose is to **make predictions** for future outcomes based on these models.

What is Machine Learning?

- The term Machine Learning was first coined by **Arthur Samuel** in the year 1959, an IBM employee and pioneer in the field of computer gaming and artificial intelligence
- Humans can learn from their experiences with their learning capability,
- But can a machine also learn from experiences or past data like a human does?
- It is the branch of AI mainly concerned with the development of algorithms which allow a computer to learn from the **data** and **past experiences** on their own.
- Machine Learning is a system that can learn from **example** through **self-improvement** and without being explicitly coded by programmer.
- Machine learning uses various algorithms for building **mathematical models** and making **predictions** using historical data or information

Traditional Programming VS Machine Learning?

Basic Difference in ML and Traditional Programming?

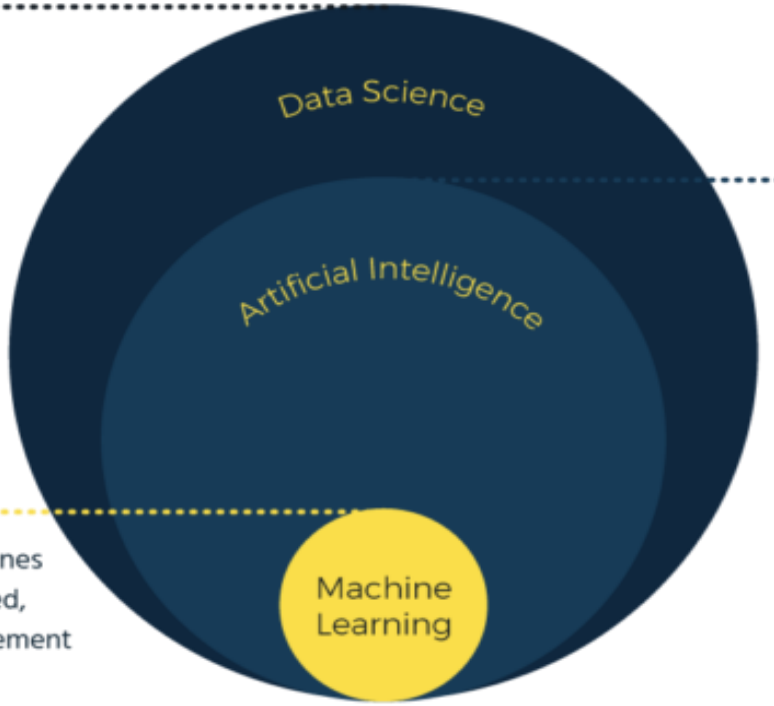
Traditional Programming : We feed in DATA (Input) + PROGRAM (logic), run it on machine and get output.

Machine Learning : We feed in DATA(Input) + Output, run it on machine during training and the machine creates its own program(logic), which can be evaluated while testing.

AI vs. Data Science vs. Machine Learning

Data Science

- Collection, preparation, and analysis of data
- Leverages AI/ML, research, industry expertise, and statistics to make business decisions



Artificial Intelligence

- Technology for machines to understand/interpret, learn, and make 'intelligent' decisions
- Includes Machine Learning among many other fields

Machine Learning

- Algorithms that help machines improve through supervised, unsupervised, and reinforcement learning
- Subset of AI and Data Science tool

Introduction to Machine Learning

innovate achieve lead

Defining the Learning Tasks

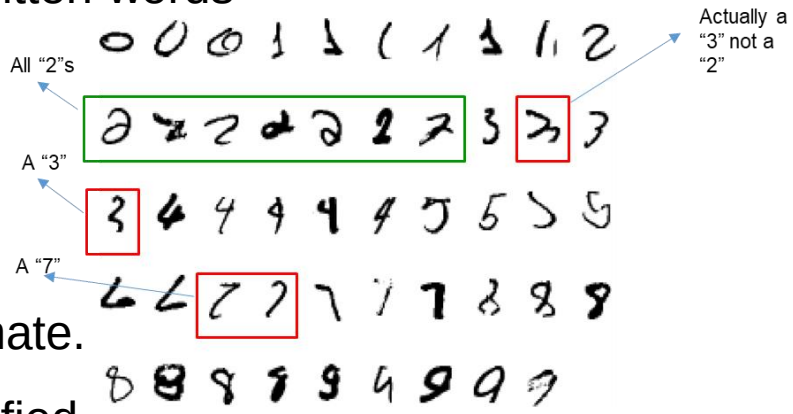
Improve on task T, with respect to performance metric P, based on experience E

Example 1

T: Recognizing hand-written words

P: Percentage of words correctly classified

E: Database of human labelled images of handwritten words



Example 2

T: Categorize email messages as spam or legitimate.

P: Percentage of email messages correctly classified.

E: Database of emails, some with human-given labels

Introduction to Machine Learning

innovate achieve lead

Defining the Learning Tasks

Improve on task T, with respect to performance metric P, based on experience E

Example 3

T: Playing Checkers

P: Percent of games won against opponents

E: Games Played against itself

Example 4

T: Drive on public four-lane highways using vision sensors.

P: Average distance travelled before an error (as judged by human).

E: A sequence of images and steering commands recorded while observing a human driver

Introduction to Machine Learning

innovate achieve lead

Applications - Perspectives

Object Categorization Prediction

Medical Diagnosis

Transaction Analysis

Recommendation System

Speech – Text Processing

Sequence

Forecasting

Medical Research

Recommendation System

Content Management

Introduction to Machine Learning

innovate achieve lead

Applications - Perspectives

Planning

Navigation

Path Finding

Gaming

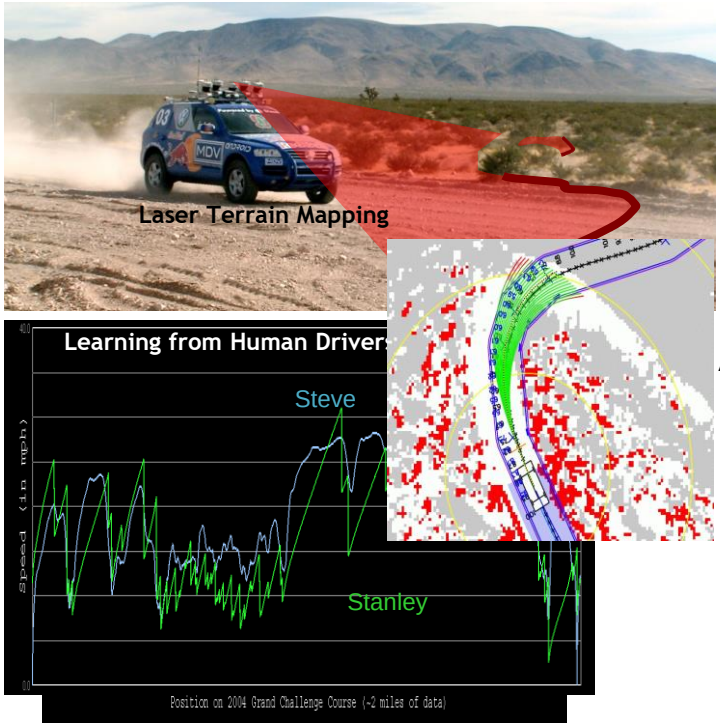
Controlling

Problem Solving

Pattern Recognition

Optimization

Decision System

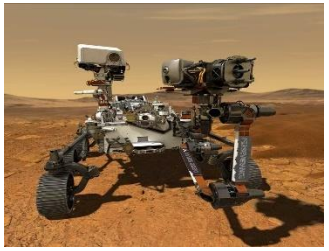


Why ML

When Do We Use Machine Learning?

ML is used when:

- Human expertise does not exist (navigating on Mars)



- Humans can't explain their expertise (Biometrics)

- Models must be customized (personalized medicine)

- Learning isn't always useful:
  - There is no need to "learn" to calculate payroll

Need For Machine Learning?

- Some tasks cannot be defined well, except by examples.
- Relationships and correlations can be hidden within large amounts of data.
- Machine - Learning may be able to find these relationships.
- The amount of knowledge available about certain tasks might be too large for explicit encoding by humans (e.g., medical diagnostic).

- We generate around 2.5 quintillion bytes **(1 Quintillion is 10^18)** of data every single day!
- **Increase in Data Generation:** Due to excessive production of data, we need a method that can be used **to structure, analyze and draw useful insights from data**. ML uses data to solve problems and find solutions to the most complex tasks faced by organizations.
- **Improve Decision Making:** By making use of various algorithms, Machine Learning can be used to make **better business decisions**. For example, Machine Learning is used to forecast sales, predict downfalls in the stock market, identify risks and anomalies, etc.

Types of ML

Introduction to Machine Learning

innovate achieve lead

Types : Inputs: Based on level of supervision

Feedback	No Feedback	Delayed Feedback (rewards/penalty)
Supervised	Unsupervised	Reinforcement

Machine Learning - Examples

Objective: Employability Prediction

Features / Attributes / Predictors

- ✓ CGPA
- ✓ Communication Skills
- ✓ Aptitude
- ✓ Programming Skills

S.No.	CGPA	Communication Skills	Aptitude	Programming Skills	Job Offered?
1	9.1	Average	Good	Excellent	Yes
2	8.4	Good	Good	Good	Yes
3	8.3	Poor	Average	Average	No
4	7.1	Average	Good	Average	No
5	8.2	Good	Excellent	Excellent	No

Machine Learning - Examples

Objective : Predicting price of a used car

Features / Attributes / Predictors

- ✓ Brand
- ✓ Year (Mfg)
- ✓ Engine Capacity
- ✓ Mileage
- ✓ Distance travelled
- ✓ Cab?

S.No	Brand	Year (Mfg)	Engine Capacity	Mileage	Distance travelled	Cab?	Price (in Rs.)
1.	Honda City ZX	2008	1100	10.5	45000	N	3,50,000
2							
3							
4							

Machine Learning - Examples

innovate achieve lead

Objective: Market Segmentation Study

Features / Attributes / Predictors

- ✓ Family income
- ✓ # of visits in a month
- ✓ Average money spent in a month
- ✓ Zip code

Customers for a retailer may fall into

- ✓ two groups say big spenders and low spenders
- ✓ three groups say big spenders, medium spenders and low spenders
- ✓ Four groups, ….

S.No.	Zip Code	Family Income	# of visits in a month	Average Money Spent in a month
1	500078	11,50,000	4	8,000

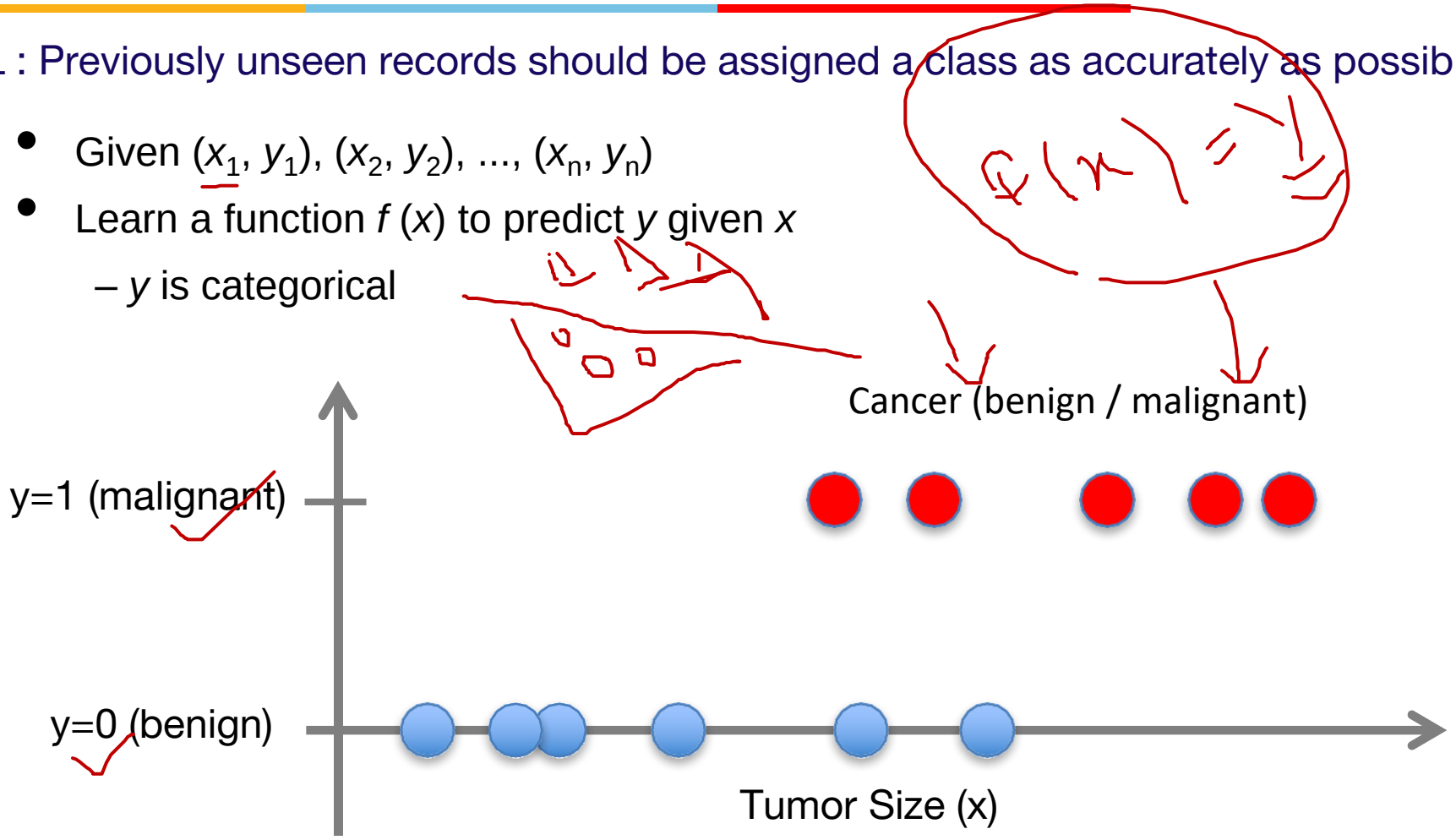
Introduction to Machine Learning

innovate achieve lead

Supervised Learning: Classification

GOAL : Previously unseen records should be assigned a class as accurately as possible.

- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is categorical

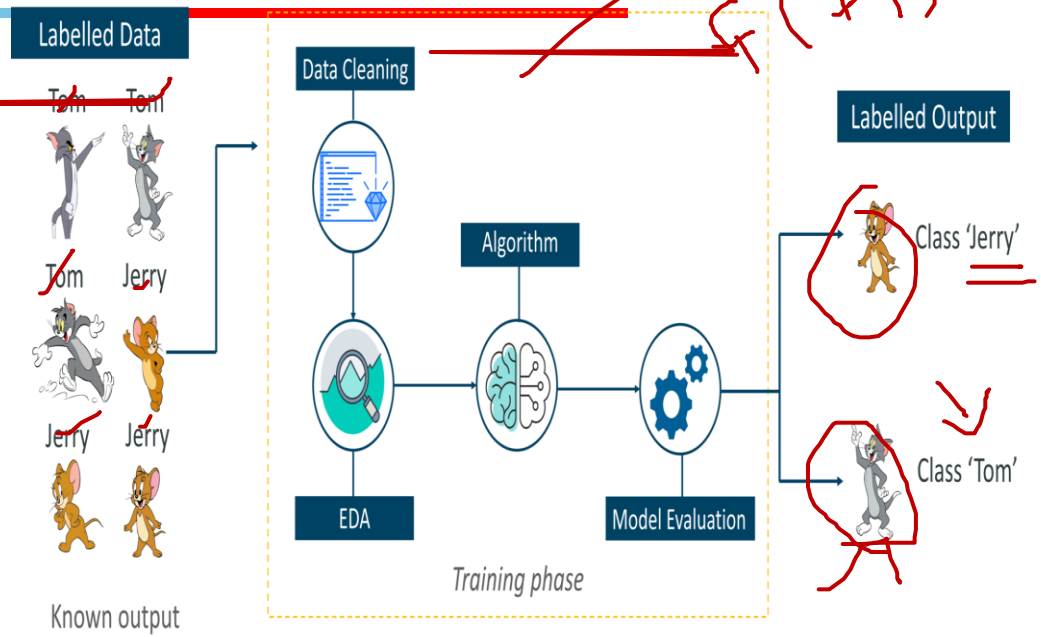


Slide Credit: Eric Eaton

Introduction to Machine Learning

Supervised Learning

- Supervised learning is a technique in which we train the machine using data which is well labeled.
- Applications in which the training data comprises examples of the input vectors along with their corresponding target vectors are known as supervised learning problems.



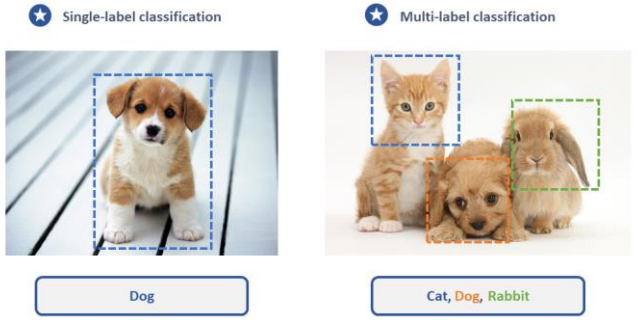
- The training data provided to the machines work as the supervisor that teaches the machines to predict the output correctly.
- It applies the same concept as a student learns in the supervision of the teacher.
- As input data is fed into the model, it adjusts its weights until the model has been fitted appropriately.
- Supervised machine learning algorithms define models that capture relationships among data.

Introduction to Machine Learning

innovate achieve lead

Classification Applications

- Google Image Classification
- Face recognition system
- Spam filters – Specific Controls
- Document tagging
- Fraud detection



Introduction to Machine Learning

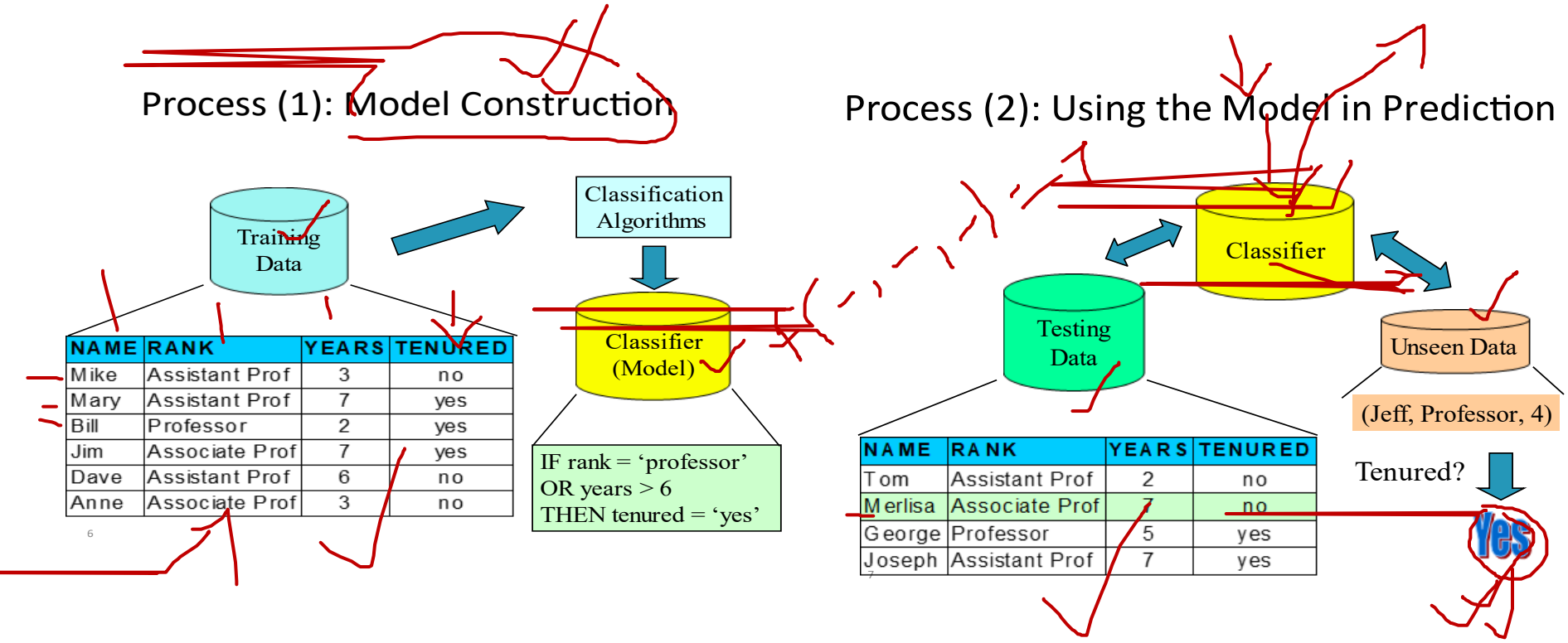
innovate achieve lead

Supervised Learning Techniques / Algorithms

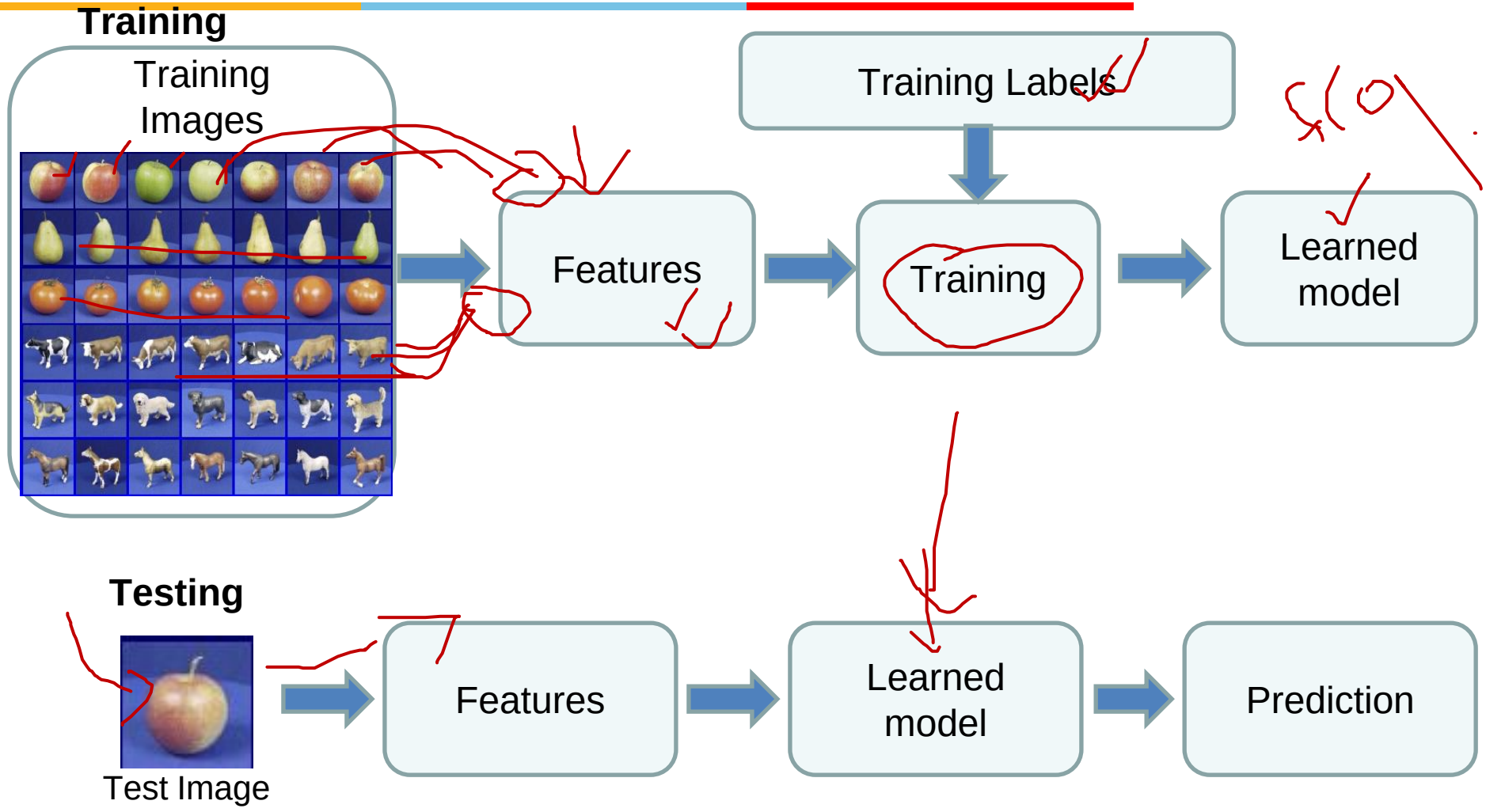
- Linear Regression
- Logistic Regression
- Naïve Bayes Classifiers
- Support Vector Machines (SVMs)
- Decision Trees and Random Forests
- Neural networks

Introduction to Machine Learning

Working of Supervised Learning Method



A Typical Supervised Learning Workflow (for Classification)



Slide credit: D. Hoiem and L. Lazebnik

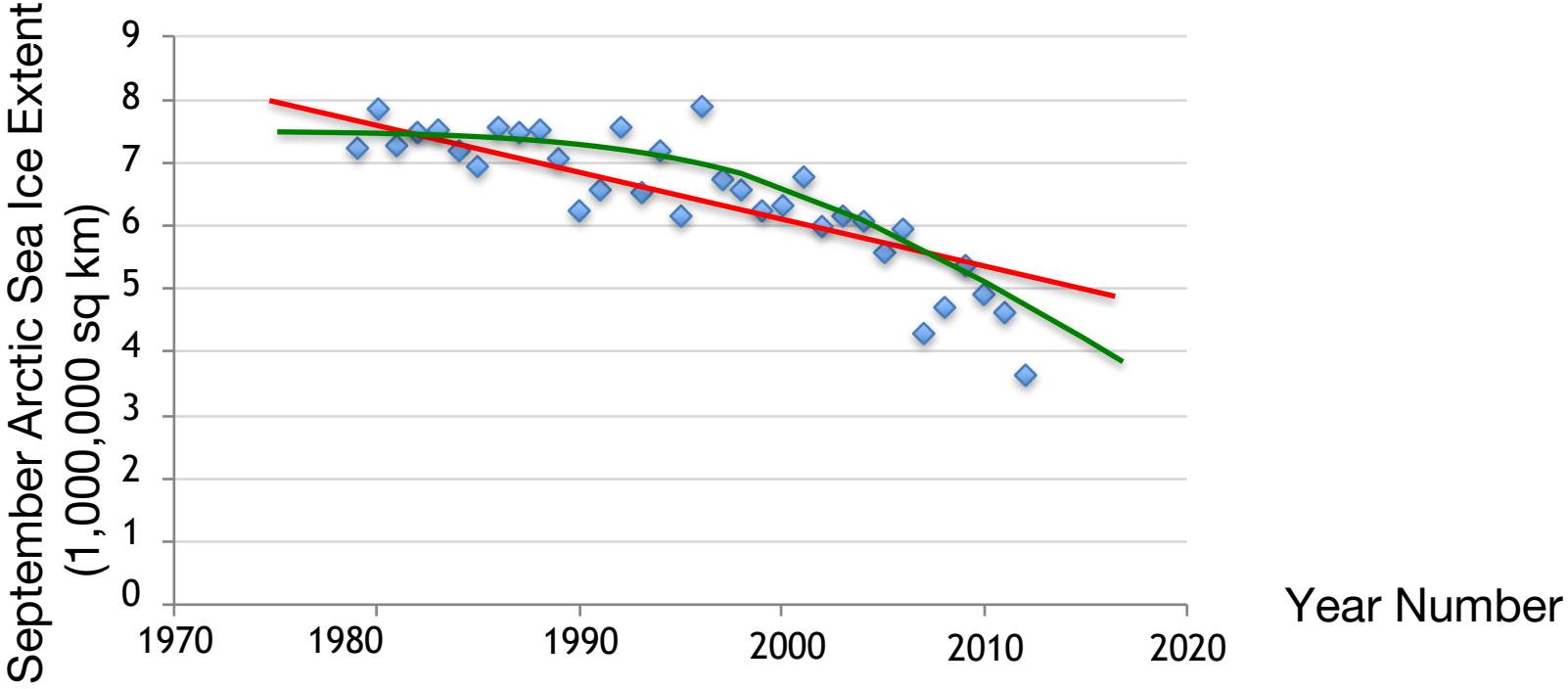
Introduction to Machine Learning

innovate achieve lead

Supervised Learning: Regression

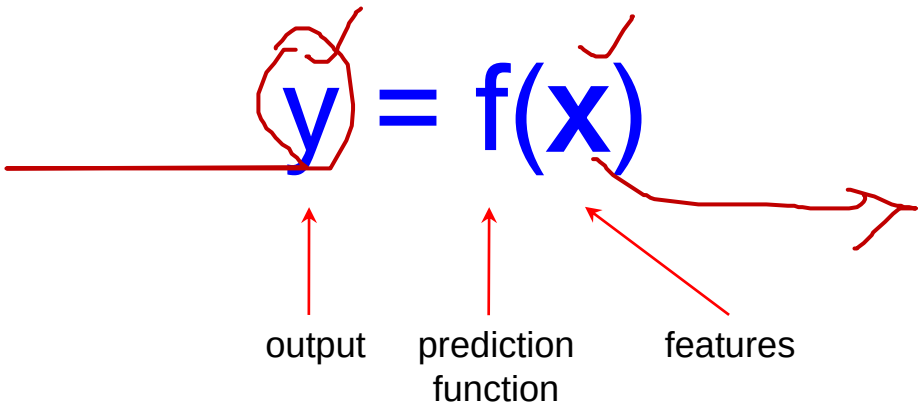
GOAL : Previously unseen records should be assigned a value as accurately as possible.

- Given (x_1, y_1), (x_2, y_2), ..., (x_n, y_n)
- Learn a function $f(x)$ to predict y given x
  - y is real-valued



Data from G. Witt. Journal of Statistics Education, Volume 21, Number 1 (2013) Slide Credit: Eric Eaton

Regression



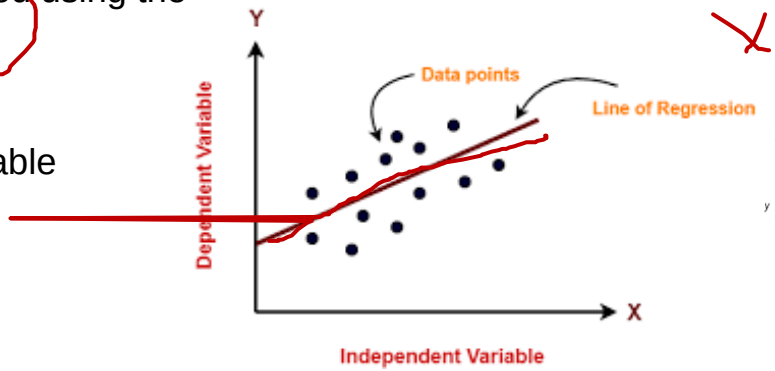
- **Training:** given a *training set* of labeled examples $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_N)\}$, estimate the prediction function f by minimizing the prediction error on the training set
- **Testing:** apply f to a never before seen *test example* $\mathbf{x}$ and output the predicted value $y = f(\mathbf{x})$

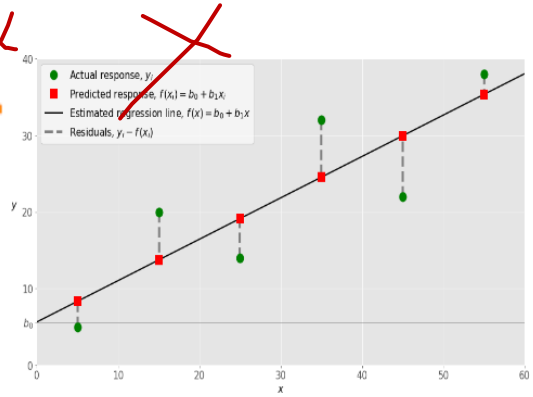
Slide credit: L. Lazebnik

Regression

- Regression analysis is used to find trends in data.
- Typically, you need regression to answer whether and how some phenomenon influences the other or how several variables are related.
- It is used for the prediction of continuous variables, such as Weather forecasting, Market Trends, etc.
- For example, you can use it to determine *if* and *to what extent* the experience or gender impact salaries.
- For example, if you've been putting on weight over the last few years, it can predict how much you'll weigh in ten years time if you continue to put on weight at the same rate
- Essentially, regression is the "best guess" at using a set of data to make some kind of prediction.
- It's fitting a set of points to a graph.

The simple linear model is expressed using the following equation: **Y = a + bX + ϵ**
Where:
Y – Dependent variable
X – Independent (explanatory) variable
a – Intercept
b – Slope
ϵ – Residual (error)





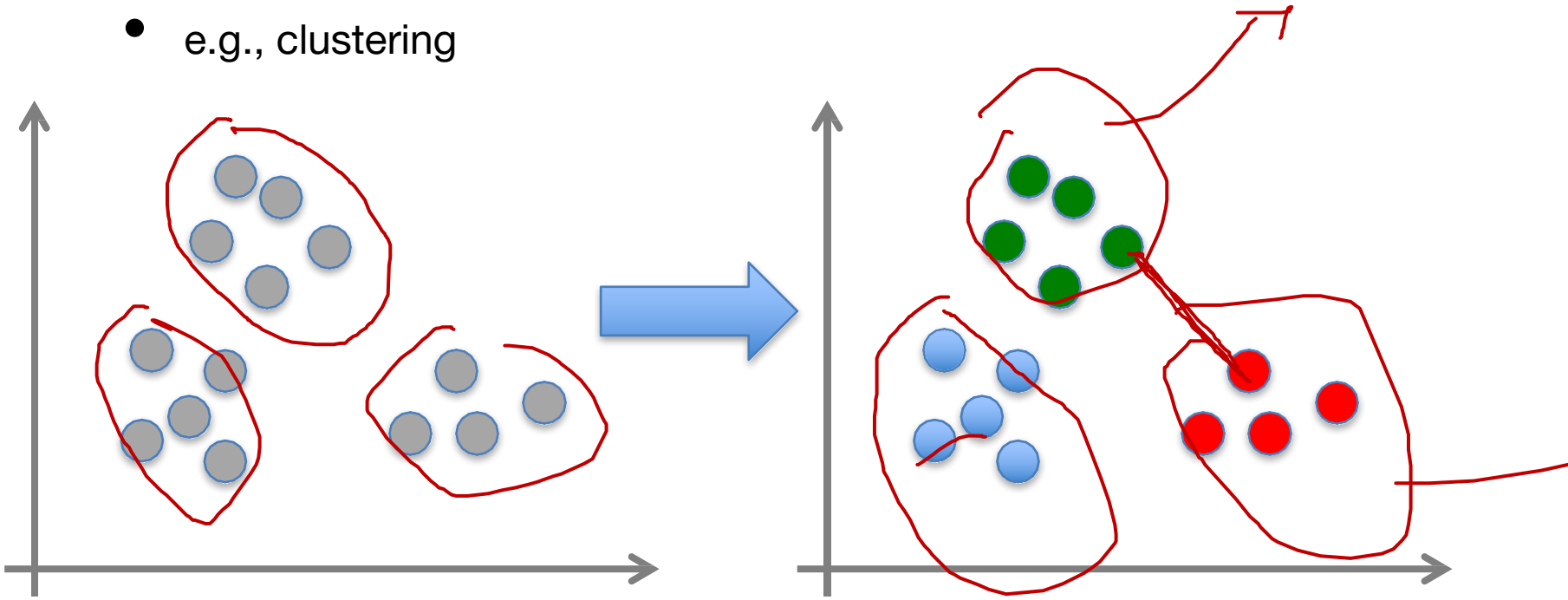
Introduction to Machine Learning

innovate achieve lead

Unsupervised Learning

GOAL : Intra cluster distances are minimized and inter cluster distances are maximized

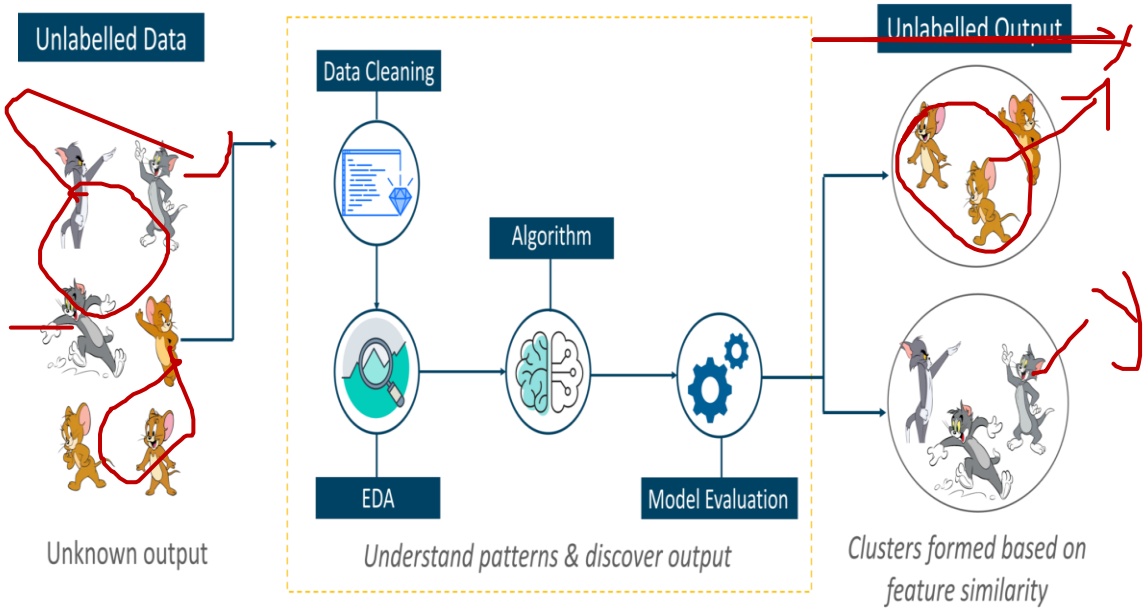
- Given x_1, x_2, ..., x_n (without labels)
- Output hidden structure behind the x's
  - e.g., clustering



Slide Credit: Eric Eaton

Introduction to Machine Learning

Unsupervised Learning



For example, it identifies prominent features of Tom such as pointy ears, bigger size, etc, to understand that this image is of type 1. Similarly, it finds such features in Jerry and knows that this image is of type 2. Therefore, it classifies the images into two different classes without knowing who Tom is or Jerry is.

Slide Credit: Eric Eaton

Introduction to Machine Learning

innovate achieve lead

Unsupervised Learning-Example

Let's, take the case of a baby and her family dog.

She knows and identifies this dog. Few weeks later a family friend brings along a dog and tries to play with the baby.

Baby has not seen this dog earlier. But it recognizes many features (2 ears, eyes, walking on 4 legs) are like her pet dog. She identifies the new animal as a dog. This is unsupervised learning, where you are not taught but you learn from the data (in this case data about a dog.)

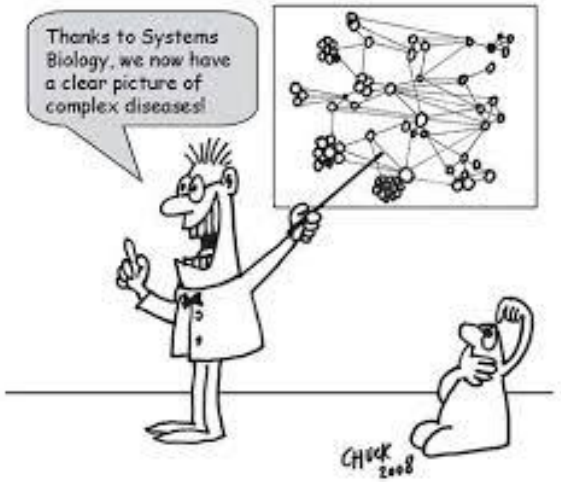
Had this been supervised learning, the family friend would have told the baby that it's a dog.

Introduction to Machine Learning

innovate achieve lead

Unsupervised Learning Applications

- Personalized recommendation system
- Targeted marketing
- Spam Filters
- Content Management – News hosted in Web
- Campaigning



Introduction to Machine Learning

innovate achieve lead

Unsupervised Learning Techniques

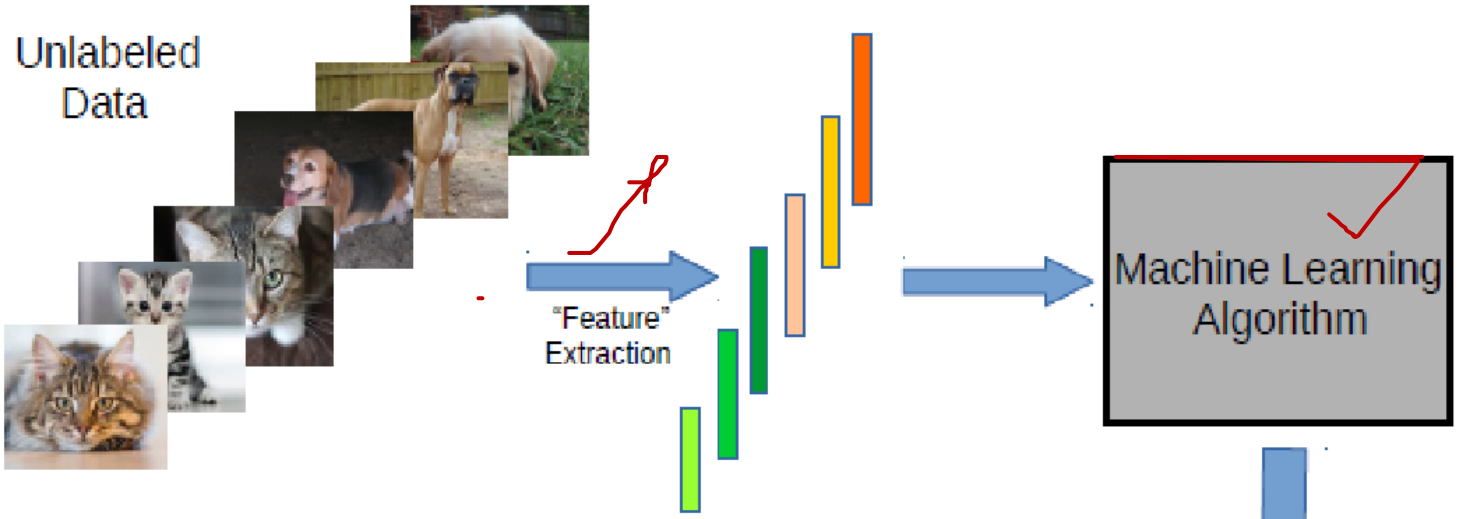
Clustering

- k-Means
- Hierarchical Cluster Analysis
- Expectation Maximization

Visualization and dimensionality reduction

- Principal Component Analysis (PCA)
- Kernel PCA
- Locally-Linear Embedding (LLE)
- t-distributed Stochastic Neighbor Embedding (t-SNE)

A Typical Unsupervised Learning Workflow (for Clustering)



Note: Unsupervised Learning too can have (and often has) a "test" phase. E.g., in this case, given a new cat/dog image, predict which of the two clusters it belongs to.

Can do it by assigning the image to the cluster with closer centroid

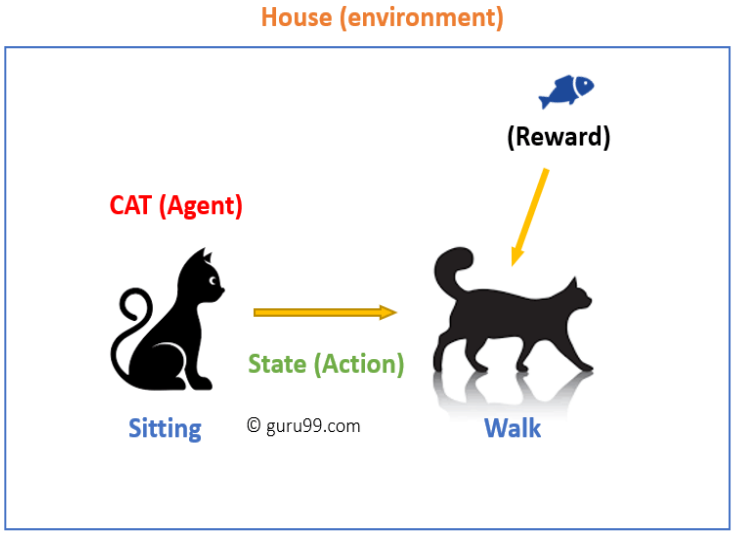
Slide Credit: Piyush Rai

Reinforcement Learning

- Reinforcement Learning is a part of Machine learning where an agent is put in an environment and he learns to behave in this environment by performing certain actions and observing the rewards which it gets from those actions.
- Imagine that you were dropped off at an isolated island! What would you do?
- As time passes by, you will learn how to live on the island. You will explore the environment, understand the climate condition, the type of food that grows there, the dangers of the island, etc. This is exactly how Reinforcement Learning works,
- it involves an **Agent** (**you, stuck on the island**) that is put in an unknown environment (island), where he must learn by observing and performing actions that result in rewards.
- Reinforcement Learning is mainly used in **advanced Machine Learning areas** such as self-driving cars, AplhaGo, etc.

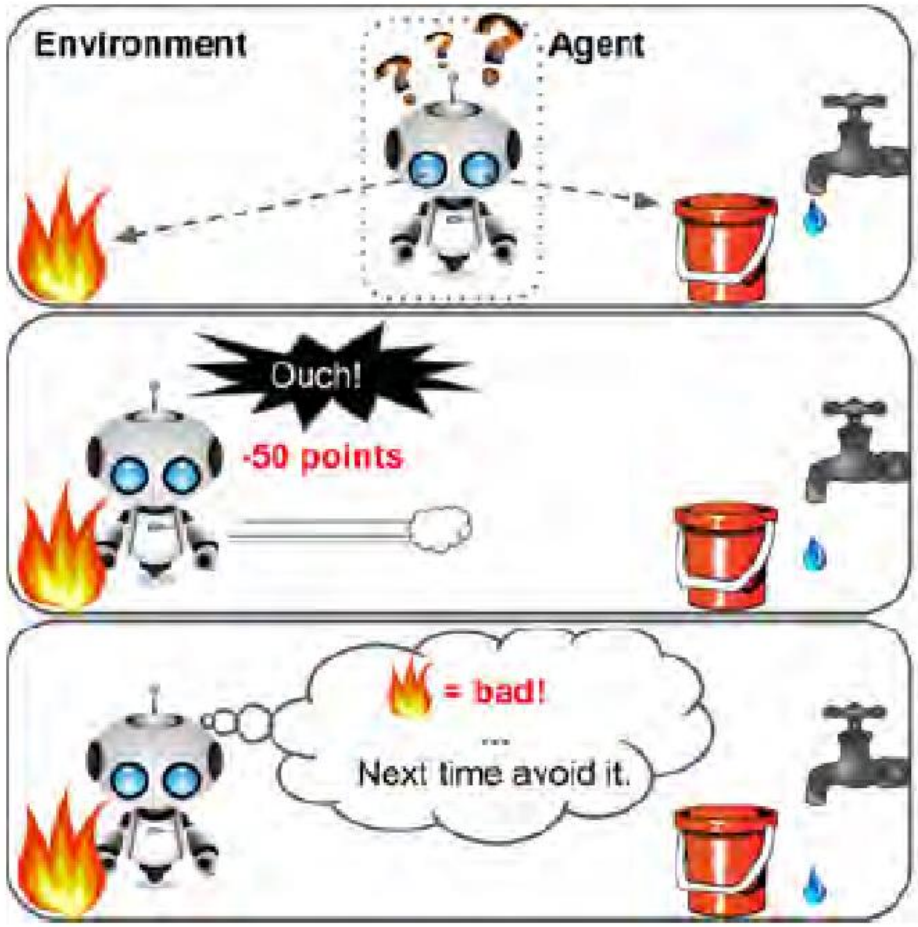
How Reinforcement Learning works?

- Consider the scenario of teaching new tricks to your cat
- As cat doesn't understand English or any other human language, we can't tell her directly what to do. Instead, we follow a different strategy.
- We emulate a situation, and the cat tries to respond in many different ways. If the cat's response is the desired way, we will give her fish.
- Now whenever the cat is exposed to the same situation, the cat executes a similar action with even more enthusiastically in expectation of getting more reward(food).
- That's like learning that cat gets from "what to do" from positive experiences.
- At the same time, the cat also learns what not do when faced with negative experiences.



- **There is no supervisor, only a real number or reward signal**
- **Sequential decision making**
- **Time plays a crucial role in Reinforcement problems**
- **Feedback is always delayed, not instantaneous**
- **Agent's actions determine the subsequent data it receives**

Reinforcement Learning



1. Observe
2. Select action using policy
3. Action!
4. Get reward or penalty
5. Update policy (learning step)
6. Iterate until an optimal policy is found

Some Examples of ML Applications

Traffic Alerts (Maps) : **Google Maps** is probably **the** app we use whenever we go out and require assistance in directions and traffic. "***Despite the Heavy Traffic, you are on the fastest route***". But, **How does it know that?**

▪Everyone using maps is providing their location, average speed, the route in which they are traveling which in turn helps Google collect massive Data about the traffic, which makes them predict the upcoming traffic and adjust your route according to it.

Social Media (Facebook) :One of the most common applications of Machine Learning is **Automatic Friend Tagging Suggestions** in Facebook or any other social media platform. Facebook uses **face detection** and **Image recognition** to automatically find the face of the person which matches it's Database and hence suggests us to tag that person based on DeepFace.

Transportation and Commuting (Uber) : If you have used an app to book a cab, you are already using Machine Learning to an extent. It provides a personalized application which is unique to you. Automatically detects your **location** and provides options to either go home or office or any other frequent place based on your **History and Patterns.**

Image courtesy: https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861

Some Examples of ML Applications

- ***Amazon's Alexa:*** The infamous Alexa, which is based on Natural Language Processing and Machine Learning is an advanced level Virtual Assistant that does more than just play songs on your playlist. It can book you an Uber, connect with the other IoT devices at home, track your health, etc.
- ***Google's Spam Filter:*** Gmail makes use of Machine Learning to filter out spam messages. It uses Machine Learning algorithms and Natural Language Processing to analyze emails in real-time and classify them as either spam or non-spam.
- ***Products Recommendations:*** Suppose you check an item on Amazon, but you do not buy it then and there. But the next day, you're watching videos on YouTube and suddenly you see an ad for the same item. You switch to Facebook, there also you see the same ad. So how does this happen?
- ***Self Driving Cars :*** Machine Learning plays a very important role in Self Driving Cars and I'm sure you guys might have heard about **Tesla**. The leader in this business and their current ***Artificial Intelligence*** is driven by hardware manufacturer **NVIDIA**, which is based on Unsupervised Learning Algorithm. they didn't train their model to detect people or any object as such. The model works on ***Deep Learning*** and it crowdsources data from all of its vehicles and its drivers.

Image courtesy: https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861

Summary : Types of Learning

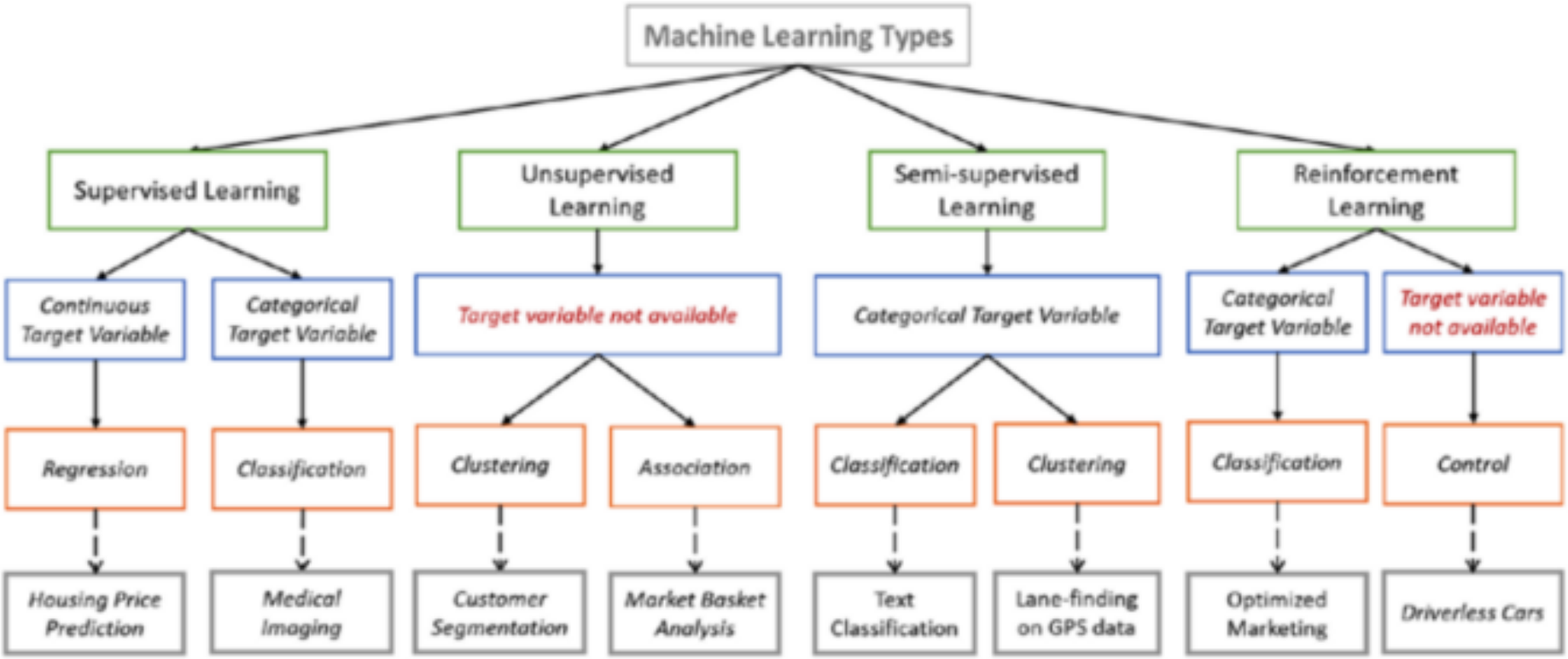


Image courtesy: https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861

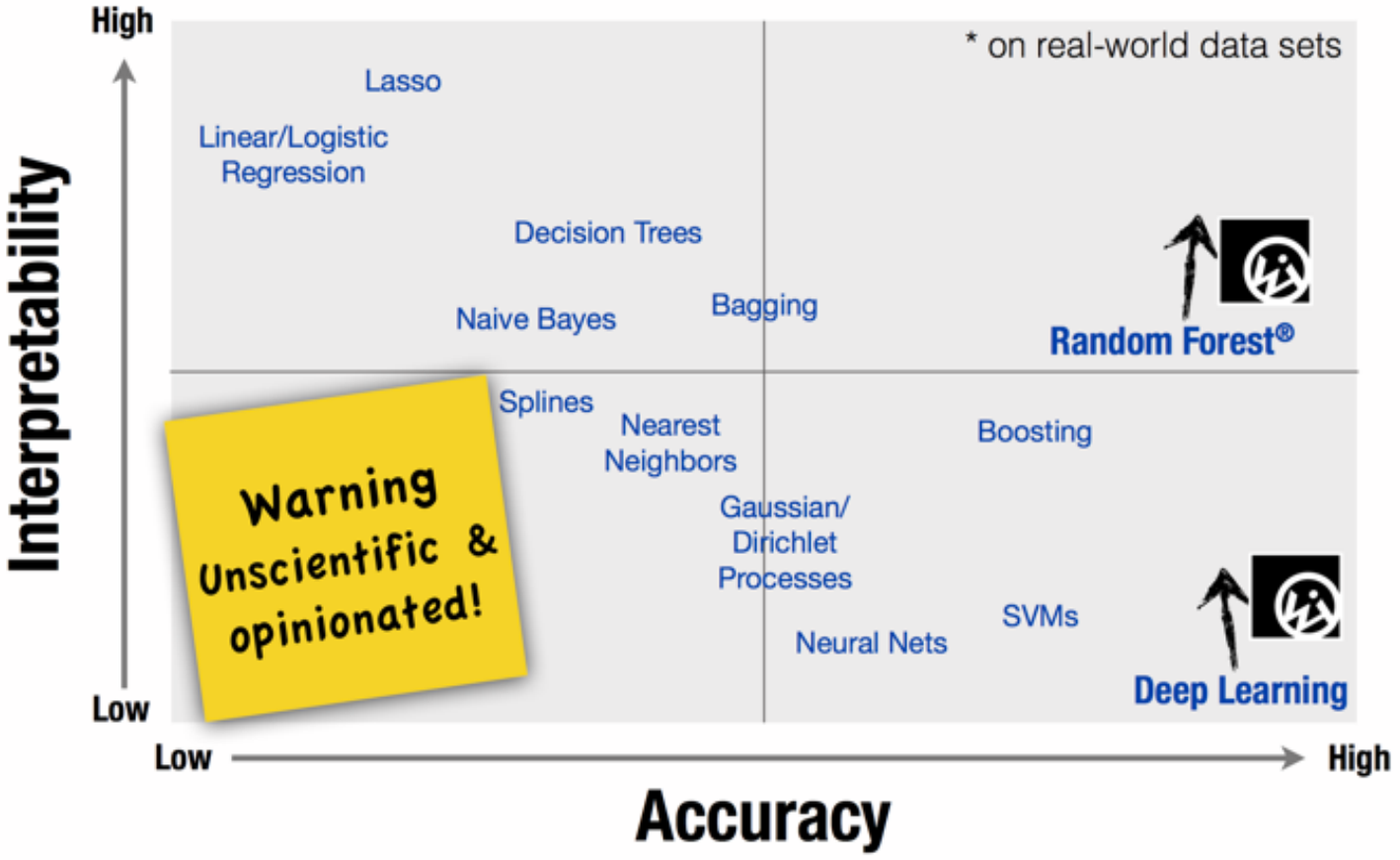
Open source ML programming tools

	Platform		Algorithms or Features
Scikit Learn	Linux, Mac OS, Windows	Python, C, C++	Classification, Regression, Clustering Preprocessing, Model Selection Dimensionality reduction.
PyTorch	Linux, Mac OS, Windows	Python, C++	Autograd Module, Optimization Module NN Module
TensorFlow	Linux, Mac OS, Windows	Python, C++	Provides a library for dataflow programming.
Weka	Linux, Mac OS, Windows	Java	Data preparation, Classification Regression, Clustering, Visualization Association rules mining

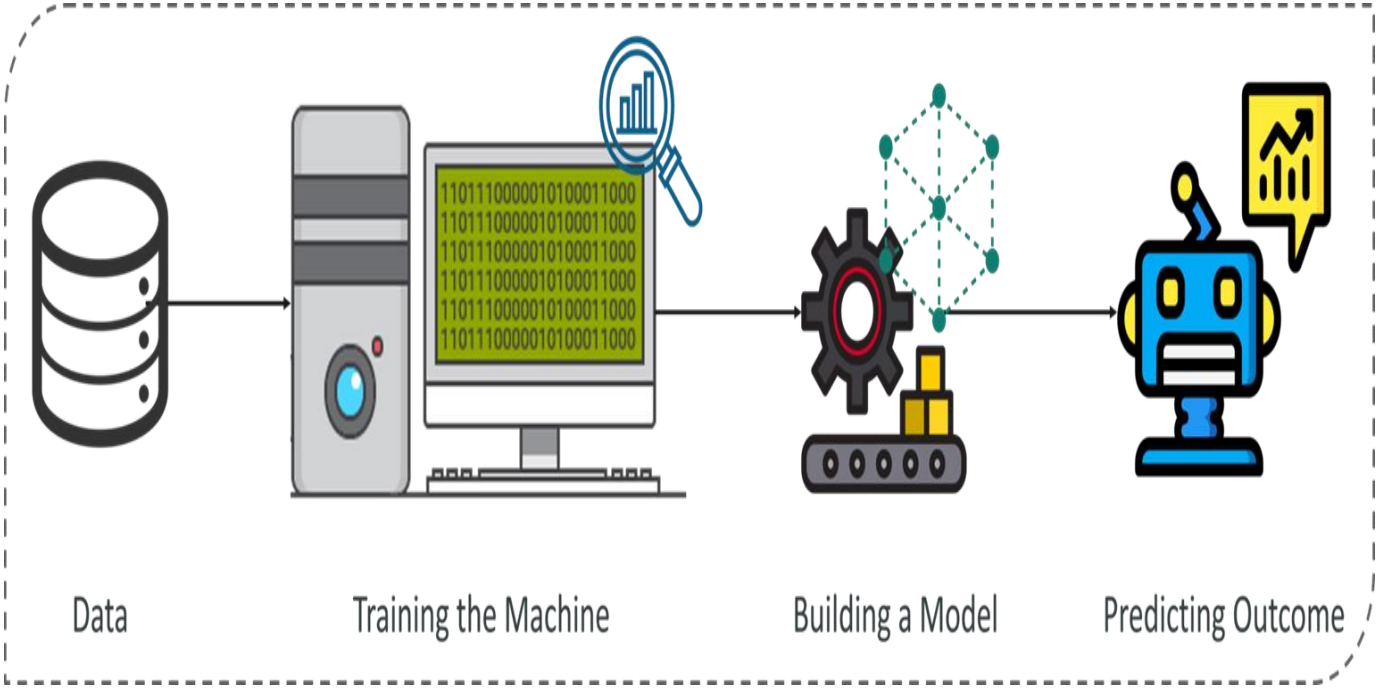
Open source ML programming tools

Colab	Cloud Service	-	Supports libraries of PyTorch, Keras, TensorFlow, and OpenCV
Apache Mahout	Cross-platform	Java Scala	Preprocessors,Regression Clustering, Recommenders Distributed Linear Algebra.
Accors.Net	Cross-platform	C#	Classification, Regression, Distribution Clustering, Hypothesis Tests & Kernel Methods, Image, Audio & Signal & Vision
Shogun	Windows Linux, UNIX Mac OS	C++	Regression, Classification, Clustering Support vector machines. Dimensionality reduction, Online learning etc.
Keras.io	Cross-platform	Python	API for neural networks

ML Algorithmic Trade-Off



ML workflow


Data
Training the Machine
Building a Model
Predicting Outcome

ML workflow

1. Should I use ML on this problem?
   - Is there a pattern to detect?
   - Can I solve it analytically?
   - Do I have data?
2. Gather and organize data.
3. Preprocessing, cleaning, visualizing.
4. Choosing a model, loss, regularization, ...
5. Optimization
6. Hyper parameter search.
7. Analyze performance and mistakes, and iterate back to step 5 (or 3)

Example: Car Price prediction based on Mileage

- Define the Objective
- Data Gathering: survey, Past Purchase data
- Data Preprocessing
  - training set; test set
  - representation of input features; output
- Exploratory Data Analysis
- Choose form of model: linear regression
- System's performance evaluation: objective function
- Optimize performance by setting appropriate parameters: Optimization
- Evaluate on test set: generalization

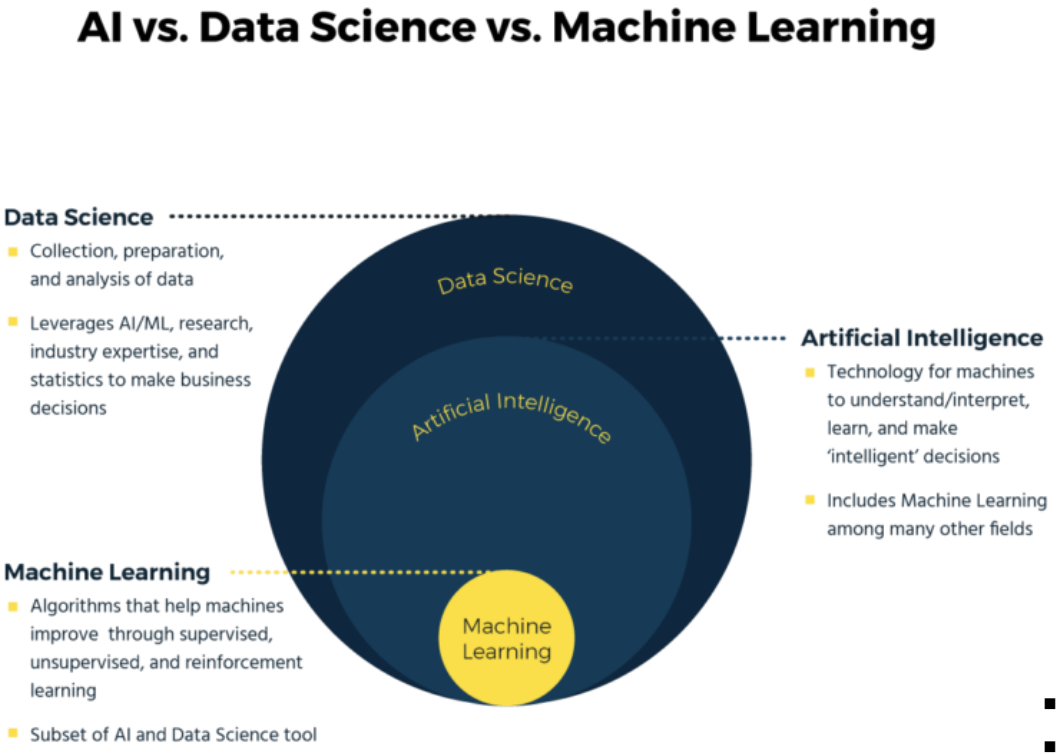
X Mileage	Y Car Price	H0	Error L1
9.8	10.48	9	2
9.12	1.75	8	7
9.5	6.95	9	3
10	2.51	9	7

Related Fields

Data Mining Vs Machine Learning

- Data mining is designed to extract the rules from large quantities of data,
- while machine learning teaches a computer how to learn and comprehend the given parameters.
- Or to put it another way, data mining is simply a method of researching to determine a particular outcome based on the total of the gathered data.
- On the other side of the coin, we have machine learning, which trains a system to perform complex tasks and uses harvested data and experience to become smarter.
- Data mining relies on vast stores of data (e.g., Big Data), which then, in turn, is used to make forecasts for businesses and other organizations.
- Machine learning, on the other hand, works with algorithms, not raw data.
- Data mining relies on human intervention and is ultimately created for use by people.
- Whereas machine learning's whole reason for existing is that it can teach itself and not depend on human influence or actions.
- Also, data mining is a process that incorporates two elements: **the database and machine learning**.
- The former provides data management techniques, while the latter supplies data analysis techniques.
- data mining can't learn or adapt, whereas that's the whole point with machine learning. Data mining follows pre-set rules and is static, while machine learning adjusts the algorithms as the right circumstances manifest themselves

Artificial Intelligence VS Data Science VS Machine Learning



Artificial Intelligence uses logic and decision trees.

Machine Learning uses statistical models.

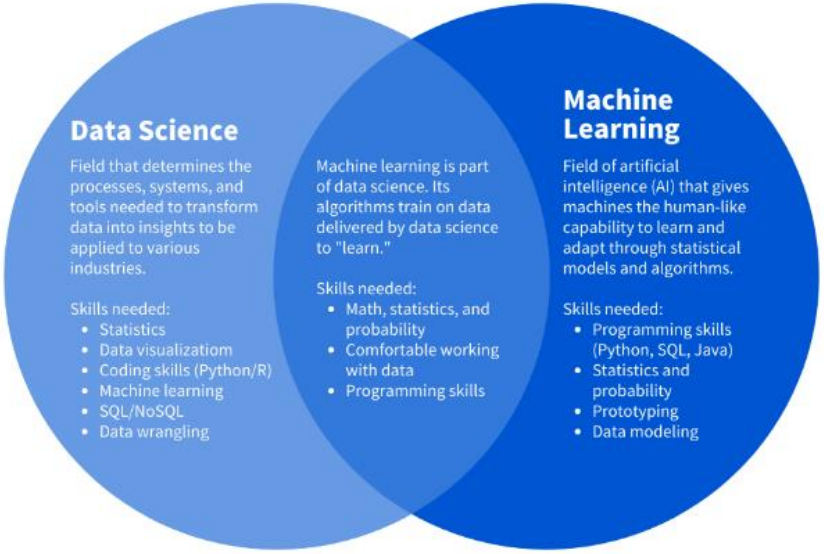
Data Science deals with structured and unstructured data.

- Data science produces insights
- Machine learning produces predictions
- Artificial intelligence produces actions

Machine Learning VS Data Science

- Data science is a field that studies data and how to extract meaning from it,
- whereas machine learning is a field devoted to understanding and building methods that utilize data to improve performance or inform predictions.
- Machine learning is a branch of artificial intelligence.

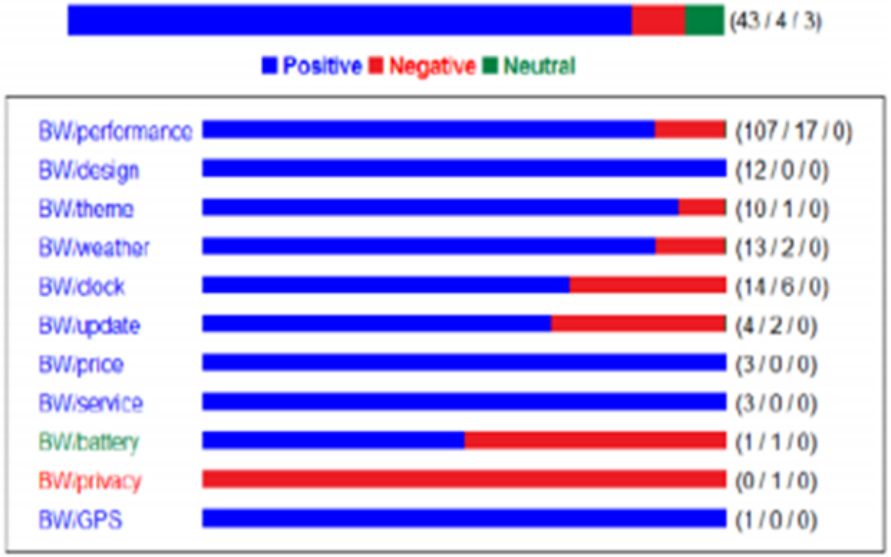
Machine learning and artificial intelligence (AI) have dominated parts of data science, playing a critical role in data analytics and business intelligence. Machine learning automates the process of data analysis and goes further to make predictions based on collecting and analyzing large amounts of data on certain populations. Models and algorithms are built to make this happen.



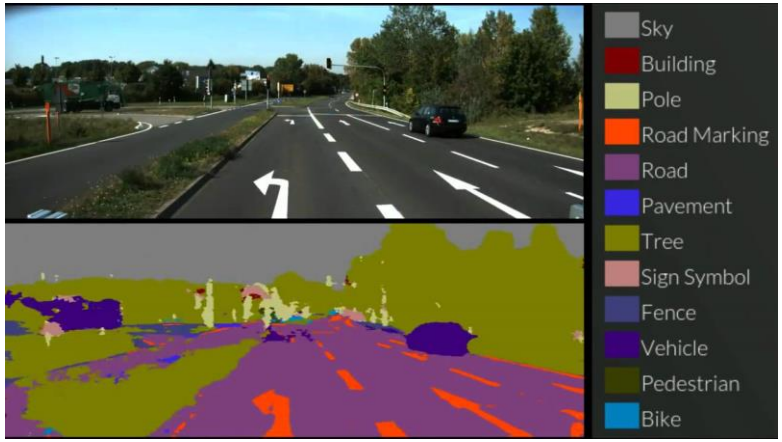
Introduction to Machine Learning

innovate achieve lead

Common Use cases - Security & Transaction Domain



Sentiment analysis on Product review of Mobile phone



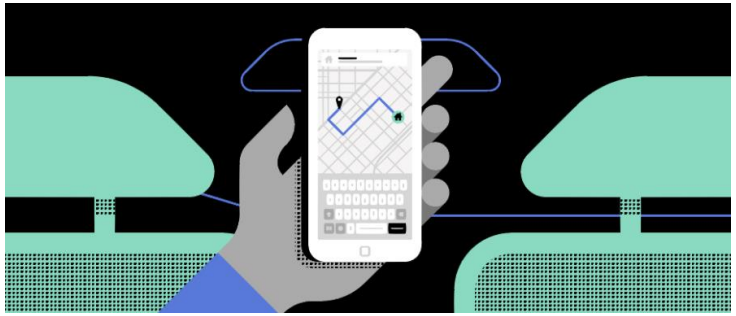
- Self Driving Cars
- Fraud Detection in Banking
- Email Filtering
- Dynamic Pricing in Travel

Derived Applications:

> Cyber Security

> Video Surveillance

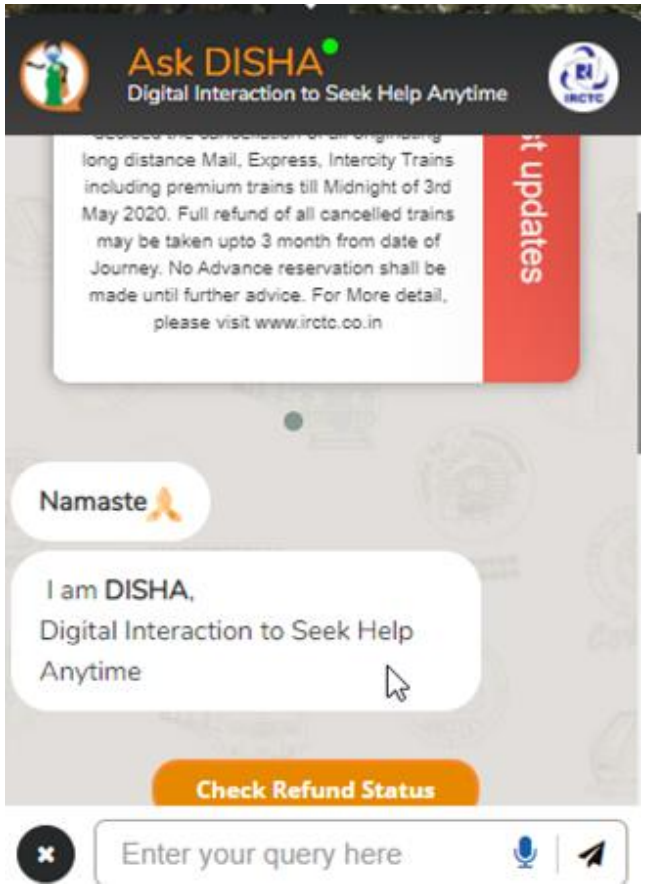
> Object Detection

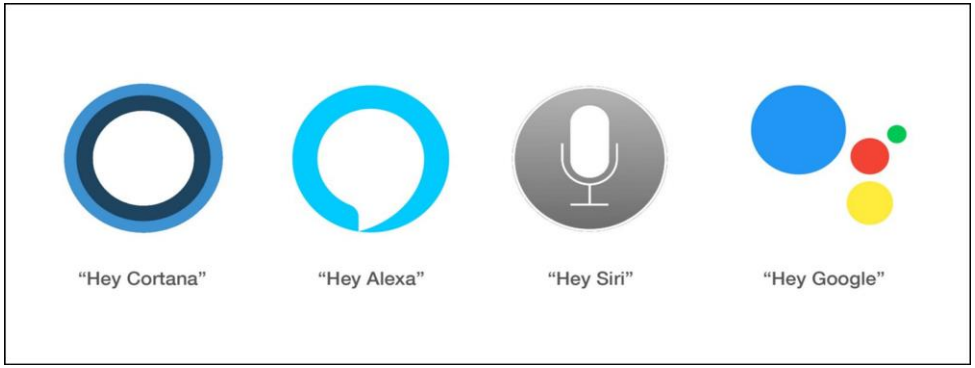


Introduction to Machine Learning

innovate achieve lead

Common Use cases - Customer Support Systems





- Apple's Siri
- Google Assistant
- Amazon's Alexa
- Google Duplex
- Microsoft's Cortana
- Samsung's Bixby

Derived Applications:

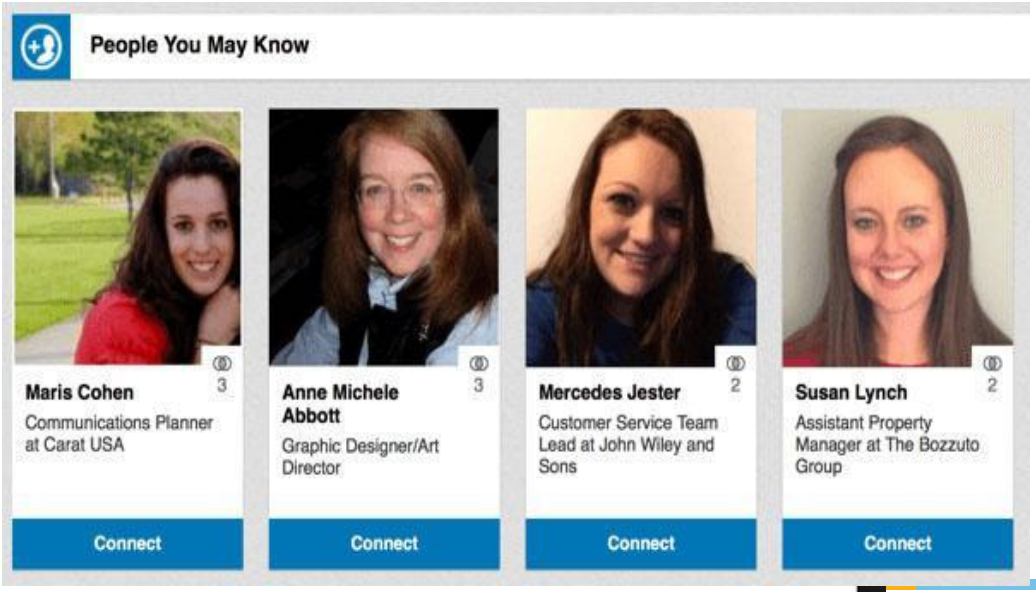
> Customer Support Query (Voice vs Text)

> Chatbots

Introduction to Machine Learning

innovate achieve lead

Common Use cases - Recommendation Engines



- E-commerce sites like Amazon and Flipkart
- Book sites like Goodreads
- Movie services like IMDb and Netflix
- Hospitality sites like MakeMyTrip, Booking.com, etc.
- Retail services like StitchFix
- Food aggregators like Zomato and Uber Eats

Derived Applications:

> Personalized Marketing

> Personalized Banking

MCQ

What is the primary goal of machine learning?
a) Minimize training time
b) Maximize model complexity
c) Minimize prediction accuracy
d) Maximize generalization to new data

Answer: d) Maximize generalization to new data

Which of the following is a type of unsupervised learning?
a) Regression
b) Classification
c) Clustering
d) Reinforcement learning

Answer: c) Clustering

What does the bias-variance tradeoff refer to in machine learning?
a) Balancing model complexity and interpretability
b) Balancing training time and prediction time
c) Balancing the model's ability to fit training data and generalize to new data
d) Balancing the tradeoff between precision and recall

Answer: c) Balancing the model's ability to fit training data and generalize to new data

MCQ

What is the main difference between supervised and unsupervised learning?
a) Supervised learning requires labeled data, while unsupervised learning does not.
b) Unsupervised learning requires labeled data, while supervised learning does not.
c) Supervised learning is used for classification, while unsupervised learning is used for regression.
d) Unsupervised learning is used for classification, while supervised learning is used for clustering.

Answer: a) Supervised learning requires labeled data, while unsupervised learning does not.

What is the purpose of a validation set in machine learning?
a) To train the model
b) To test the model
c) To fine-tune hyperparameters and avoid overfitting
d) To evaluate the model's performance on unseen data

Answer: c) To fine-tune hyperparameters and avoid overfitting

Which of the following is an example of a classification problem?
a) Predicting house prices
b) Identifying spam emails
c) Estimating stock prices
d) Forecasting temperature

Answer: b) Identifying spam emails

References

- Chapter 1 – Machine Learning, Tom Mitchell
- Chapter 1, 2 – Introduction to Machine Learning, 2nd edition, Ethem Alpaydin
- Chapter 1 - Pattern Recognition & Machine Learning Christopher M. Bhishop
- http://www.cs.princeton.edu/courses/archive/spr08/cos511/ [Web]
- https://www.softwaretestinghelp.com/machine-learning-tools/

Thank you !