Machine Learning

DSECL ZG565

M1 : Introduction

Course Faculty of MTech Cluster

BITS - CSIS

BITS Pilani

Pilani Campus

Machine Learning

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

- **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	Linear Models for Regression
M3	Linear Models for Classification
M4	Decision Tree
M5	Instance Based Learning
M6	Support Vector Machine
M7	Bayesian Learning
M8	Ensemble Learning
M9	Unsupervised Learning
M10	Machine Learning Model Evaluation/Comparison

Text books and Reference book(s)

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 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

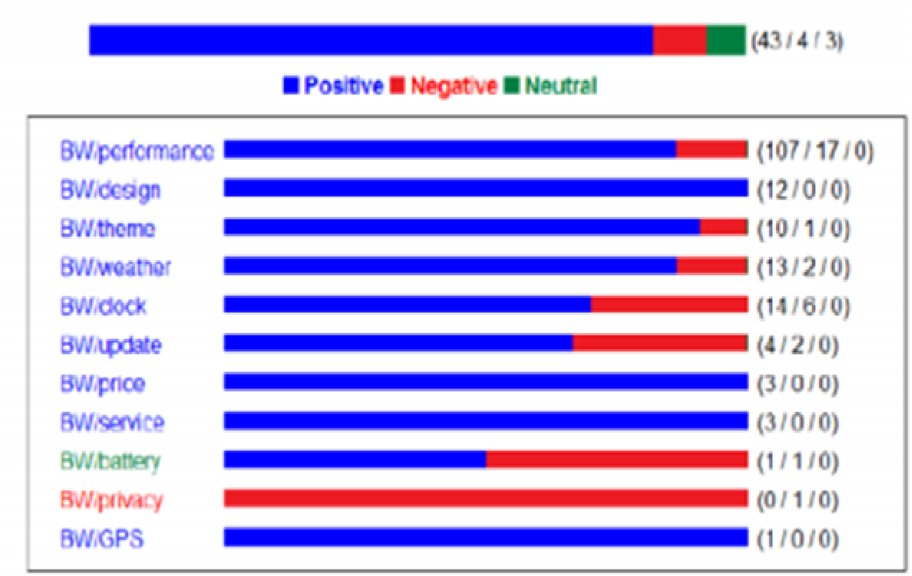
- What is Machine Learning?
- Why Machine Learning is important?
- Types of Machine Learning
- Application Areas
- Issues in Machine Learning
- Demo Case study

Common Use Cases

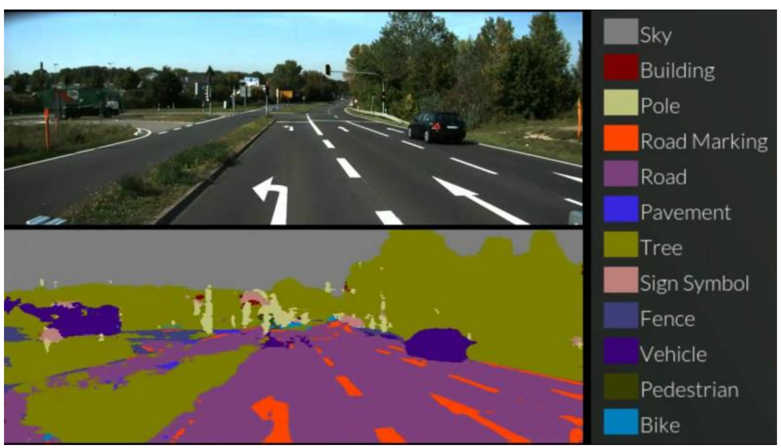
1. Applications in Smartphones
2. Transportation Optimization
3. Popular Web Services
4. Sales and Marketing
5. Security
6. Financial Domain

Introduction to Machine Learning

Common Use cases - Security & Transaction Domain



Sentiment analysis on Product review of Mobile phone



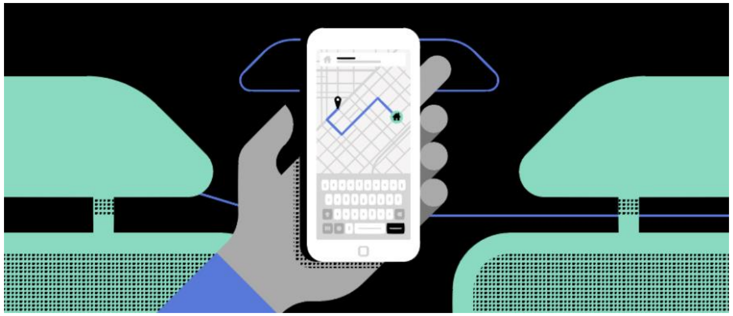
- Self Driving Cars
- Fraud Detection in Banking
- Email Filtering
- Dynamic Pricing in Travel

Derived Applications:

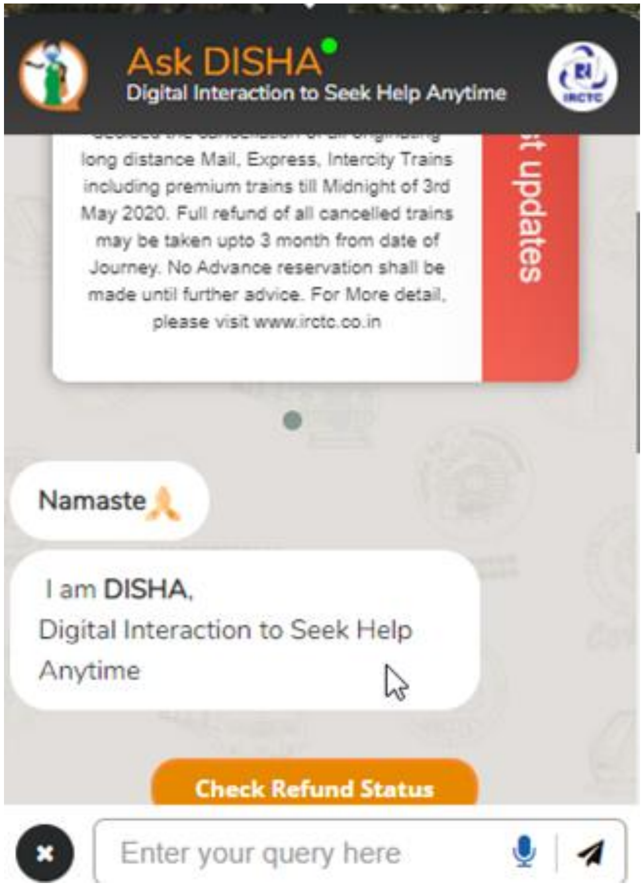
> Cyber Security

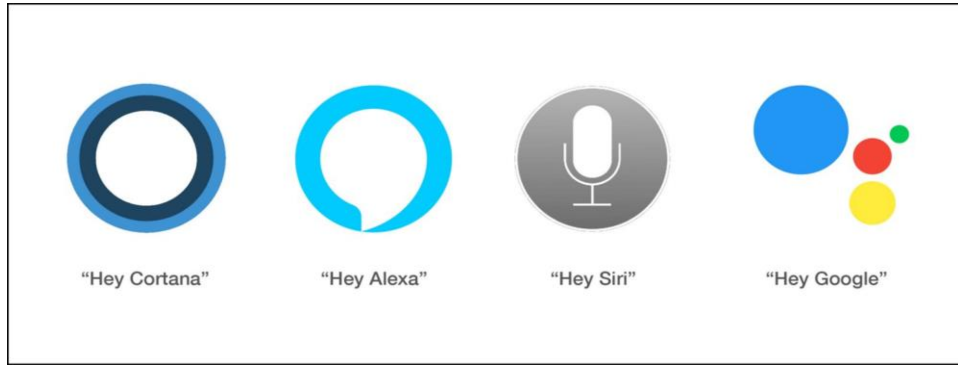
> Video Surveillance

> Object Detection



Common Use cases - Customer Support Systems





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

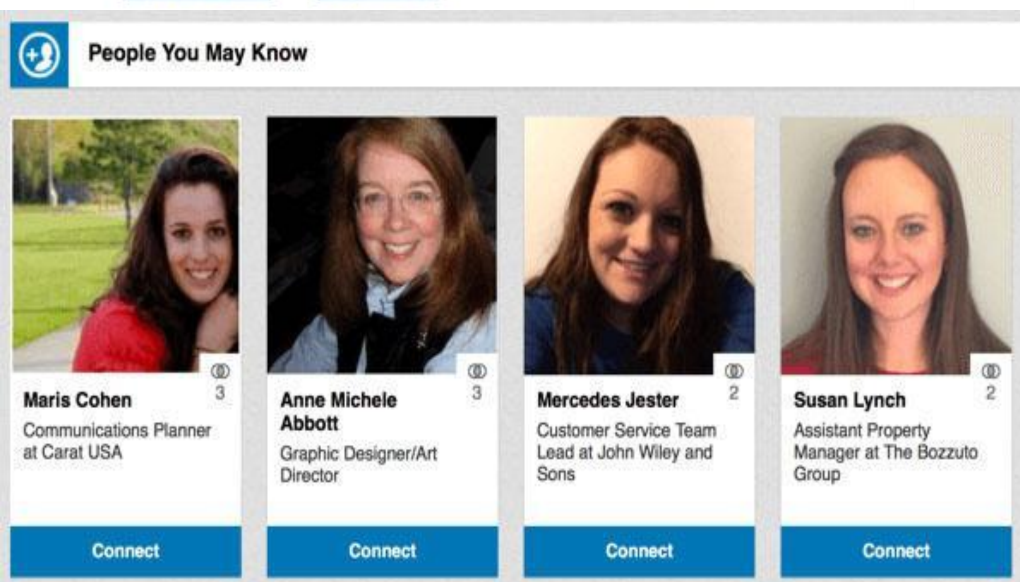
Derived Applications:

> Customer Support Query (Voice vs Text)

> Chatbots

Introduction to Machine Learning

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

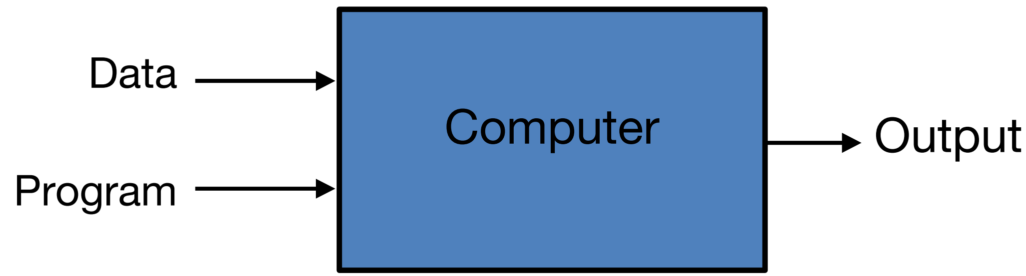
ML – What, When, Where ?

Introduction to Machine Learning

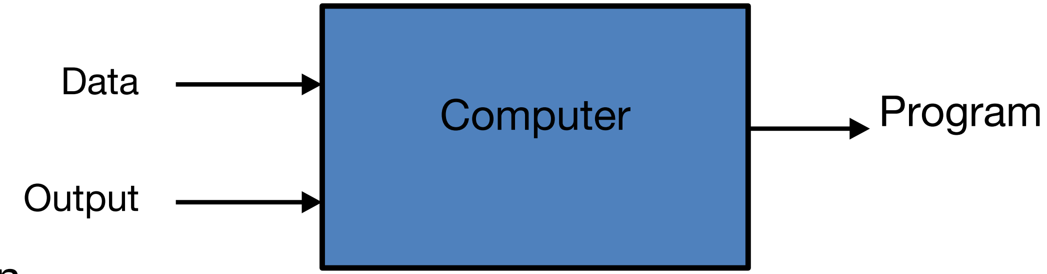
What is Machine Learning (ML)?

- The science (and art) of programming computers so they can *learn from data*
- More general definition
  - Field of study that gives computers the ability to learn without being explicitly programmed
- Engineering-oriented definition
  - Algorithms that improve their performance P at some task T with experience E
  - A well-defined learning task is given by <*P*, *T*, *E*>

Traditional programming



Machine Learning



Slide credit: Pedro Domingos

Defining the Learning Tasks

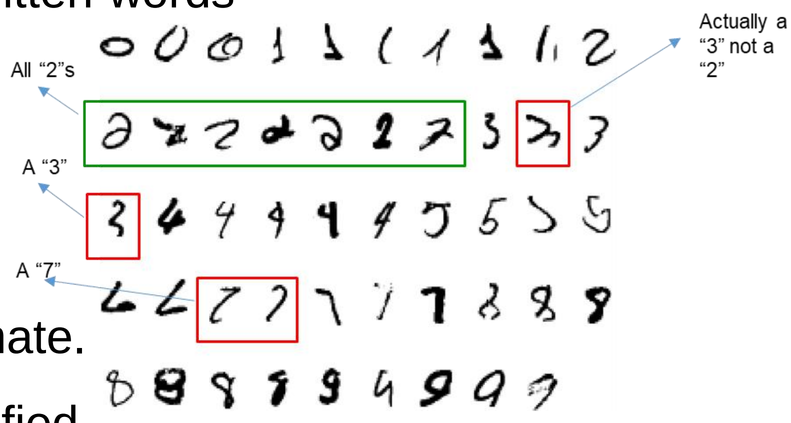
Improve on task T, with respect to performance metric P, based on experience E

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

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

Traditional Approach - Spam Filtering

Spam typically uses words or phrases such as “4U,” “credit card,” “free,” and “amazing”

Solution

Write a detection algorithm for frequently appearing patterns in spams

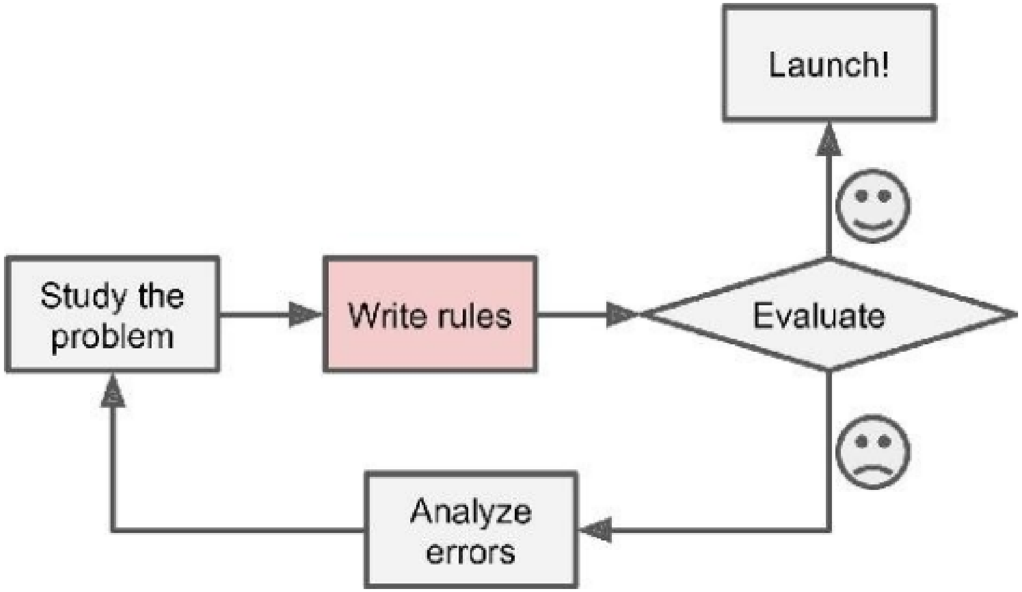
Test and update the detection rules until it is good enough.

Challenge

Detection algorithm likely

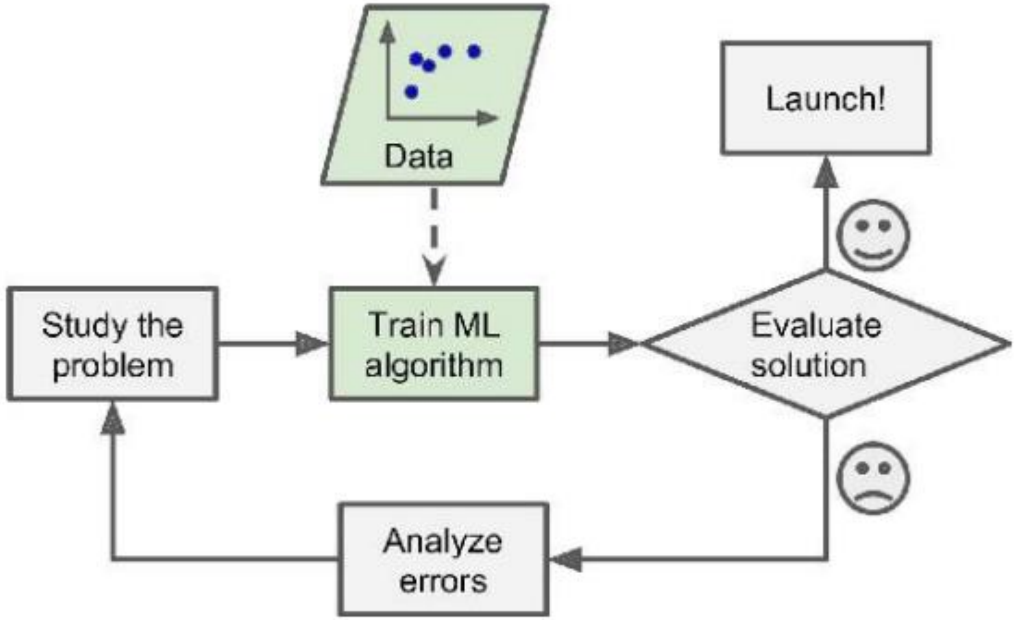
to be a long list of complex rules

hard to maintain.



ML Approach - Spam Filtering

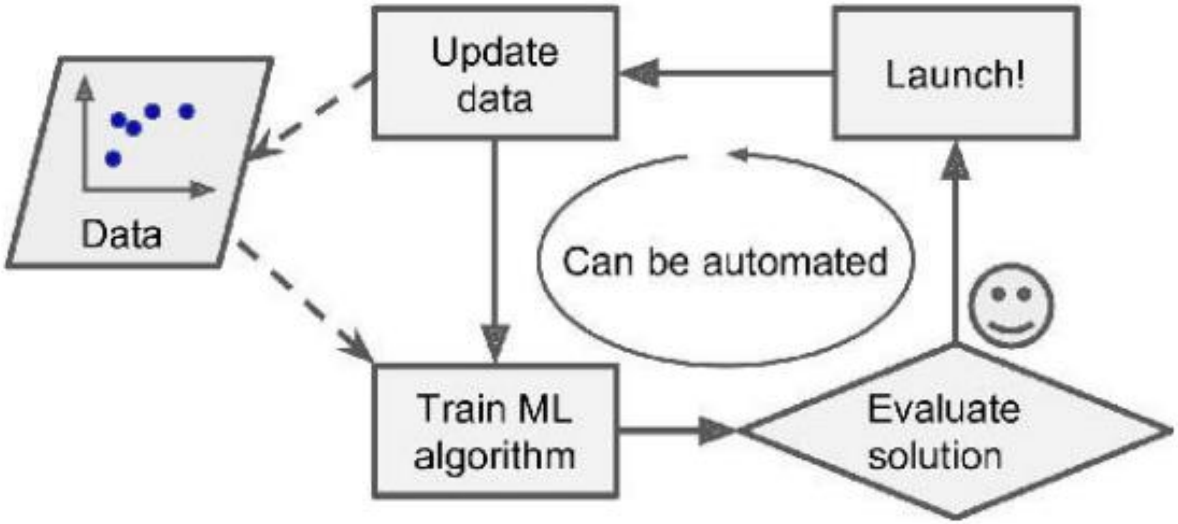
Automatically learns phrases that are good predictors of spam by detecting unusually frequent patterns of words in spams compared to “ham”s



The program is much shorter, easier to maintain, and most likely more accurate.

ML Approach - Spam Filtering

Automatically learns phrases that are good predictors of spam by detecting unusually frequent patterns of words in spams compared to “ham”s



The program is much shorter, easier to maintain, and most likely more accurate.

Introduction to Machine Learning

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

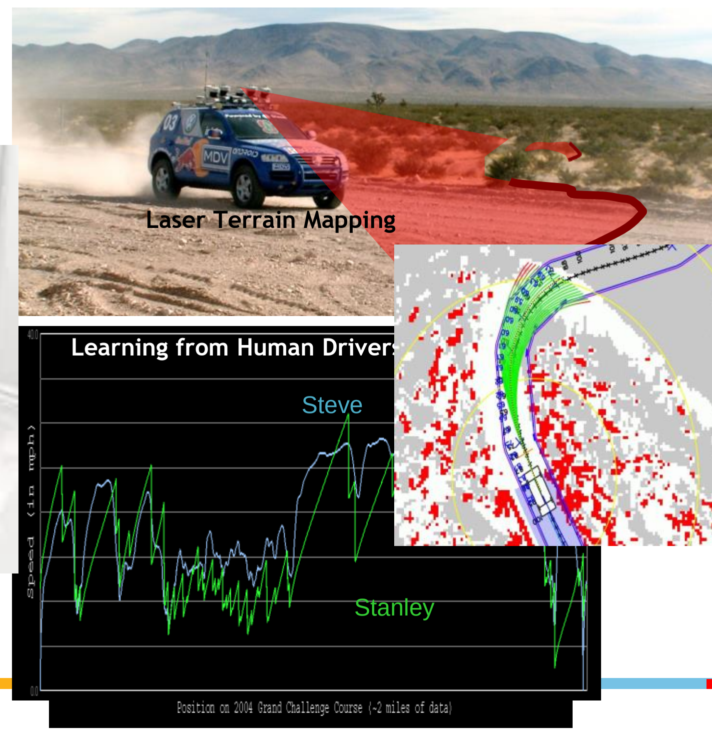
Applications - Perspectives

Planning

- Navigation
- Path Finding
- Gaming
- Controlling

Problem Solving

- Pattern Recognition
- Optimization
- Decision System



Adaptive Vision

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)

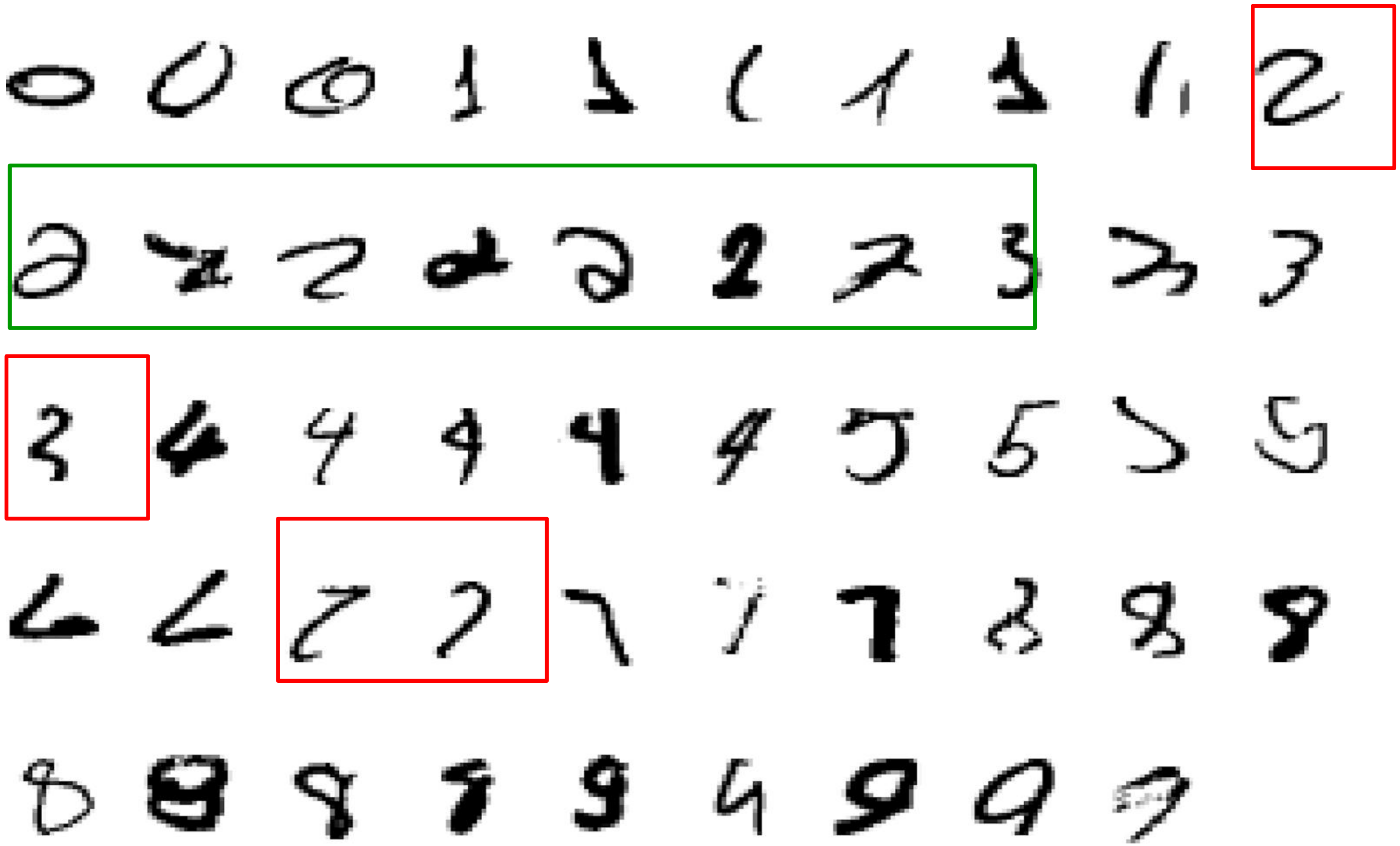
- Learning isn’t always useful:
    - There is no need to “learn” to calculate payroll

Why only ML?

- Some tasks cannot be defined well, except by examples.
  - It is very hard to write programs that solve problems like recognizing a handwritten digit
  - What distinguishes a 2 from a 7?
  - How does our brain do it
- Hidden relationships and correlations in data
- large data makes it difficult for explicit encoding by humans (e.g., medical diagnostic)
- Continuous availability of new knowledge

Pattern recognition

It is very hard to say what makes a 2



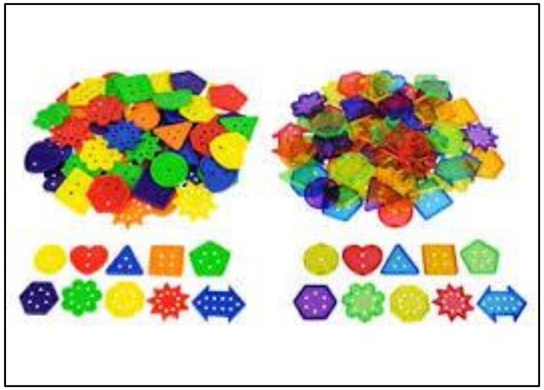
Slide credit: Geoffrey Hinton

Types of ML

Introduction to Machine Learning

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

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

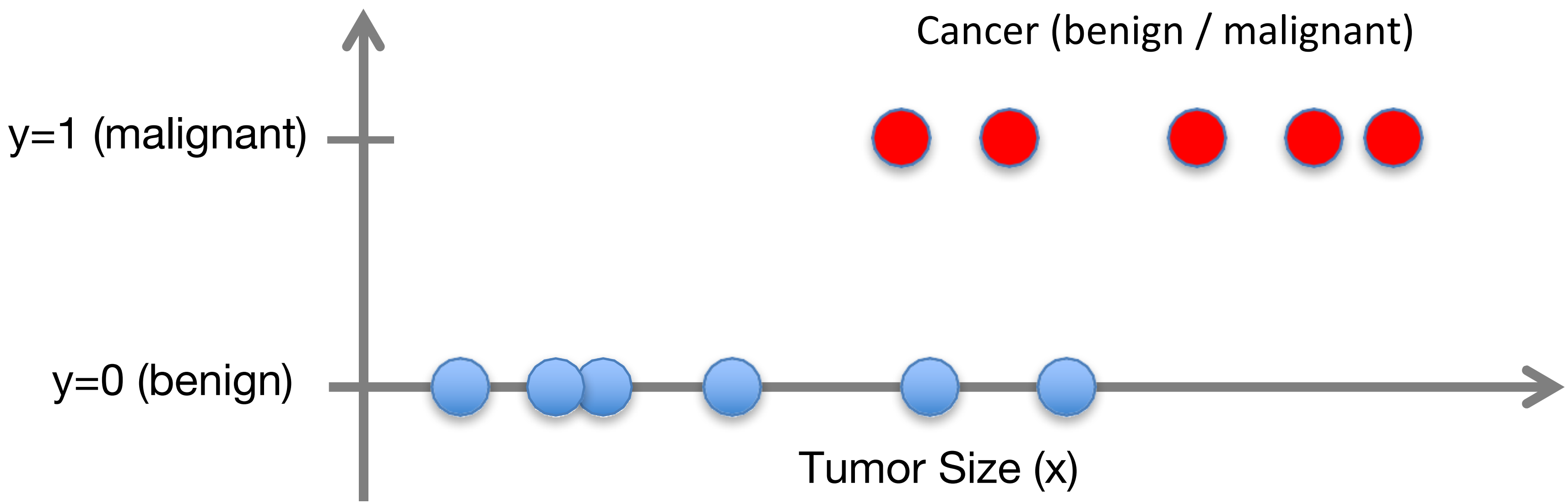
- ✓ two groups say big spenders and low spenders
- ✓ three groups say big spenders, medium spenders and low spenders
- ✓ Four groups, ….

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

Supervised Learning: Classification

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

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

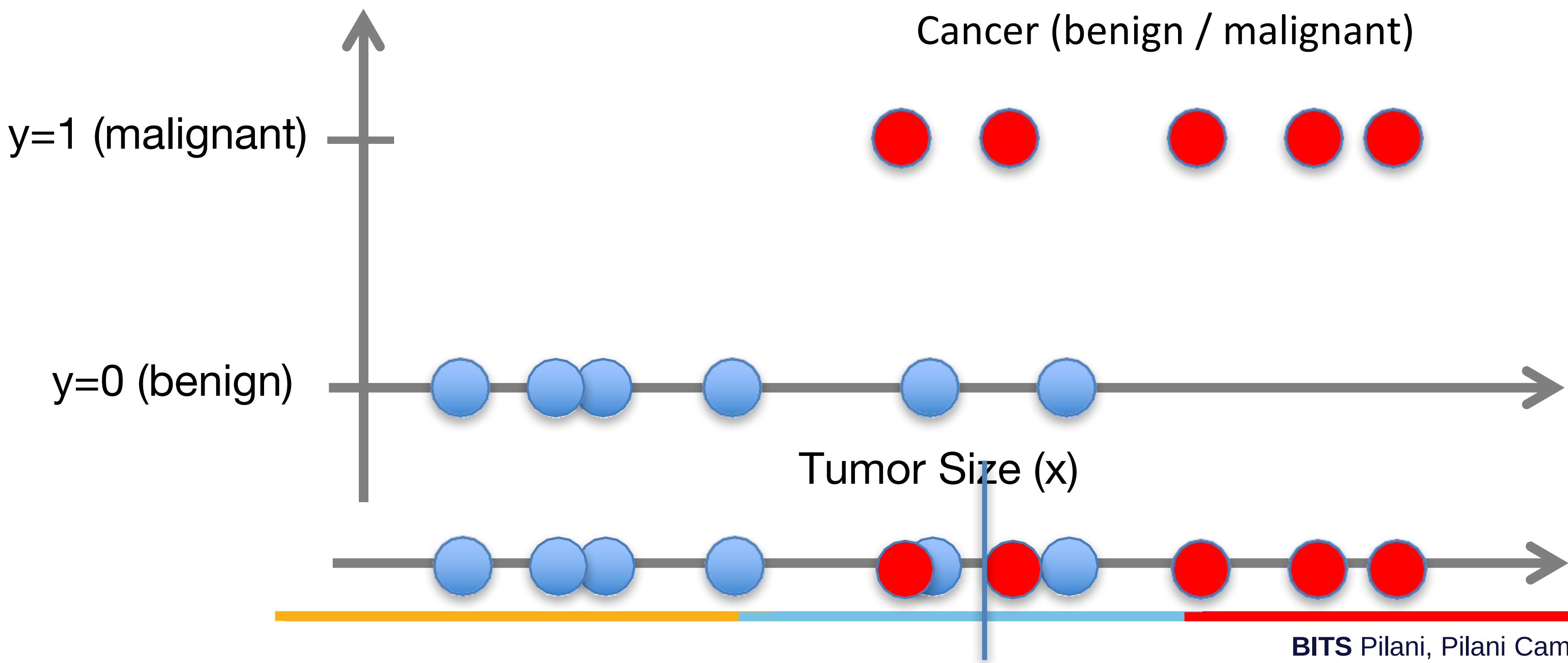


Slide Credit: Eric Eaton

Supervised Learning: Classification

- 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 categorical

Learnt classifier
If x>T, malignant else benign

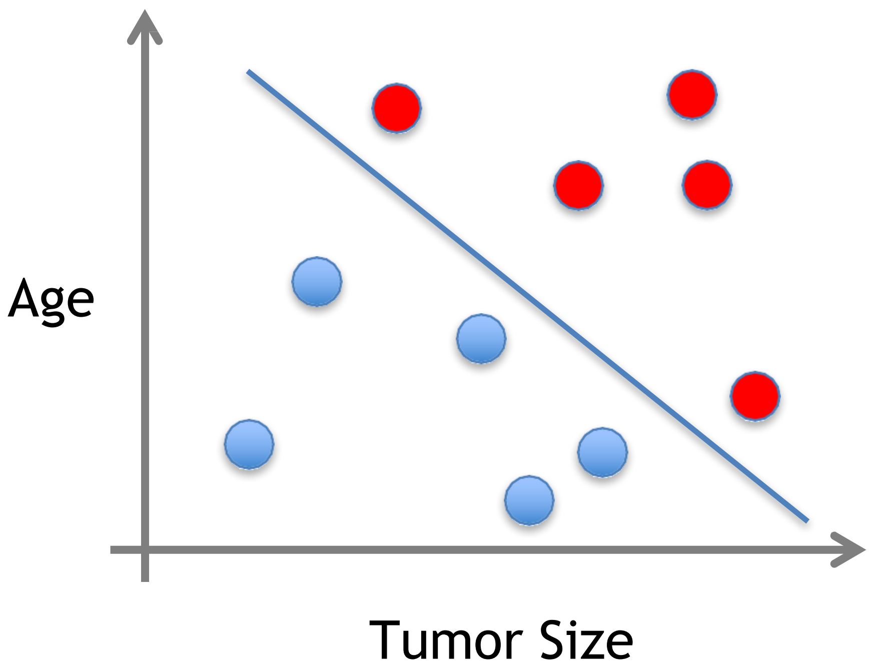


Classification Applications

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

Supervised Learning: Classification

- *x* can be multi-dimensional
  - Each dimension corresponds to an attribute



Increasing Feature Dimension

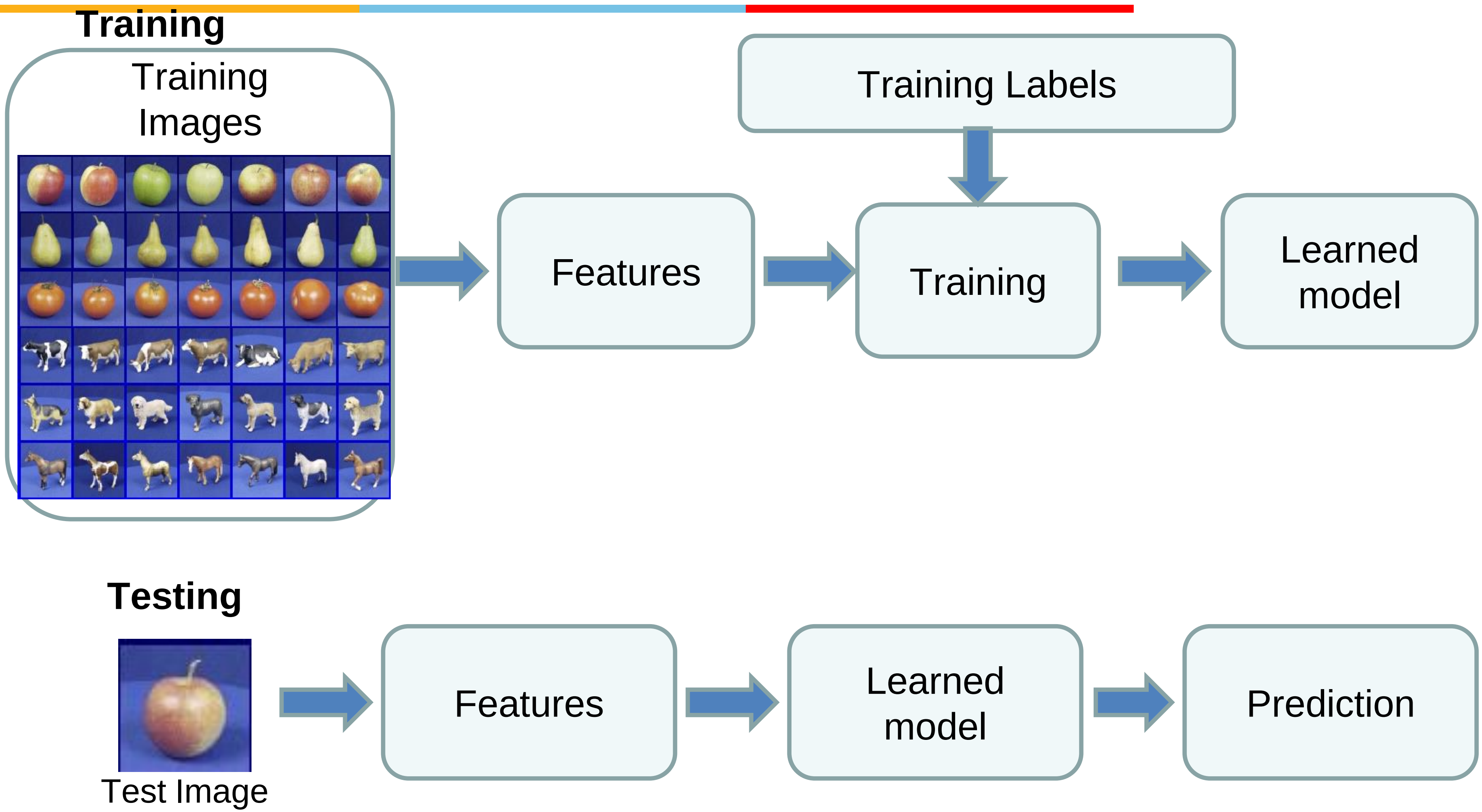
- Clump Thickness
- Uniformity of Cell Size
- Uniformity of Cell Shape

…

Supervised Learning Techniques / Algorithms

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

A Typical Supervised Learning Workflow (for Classification)

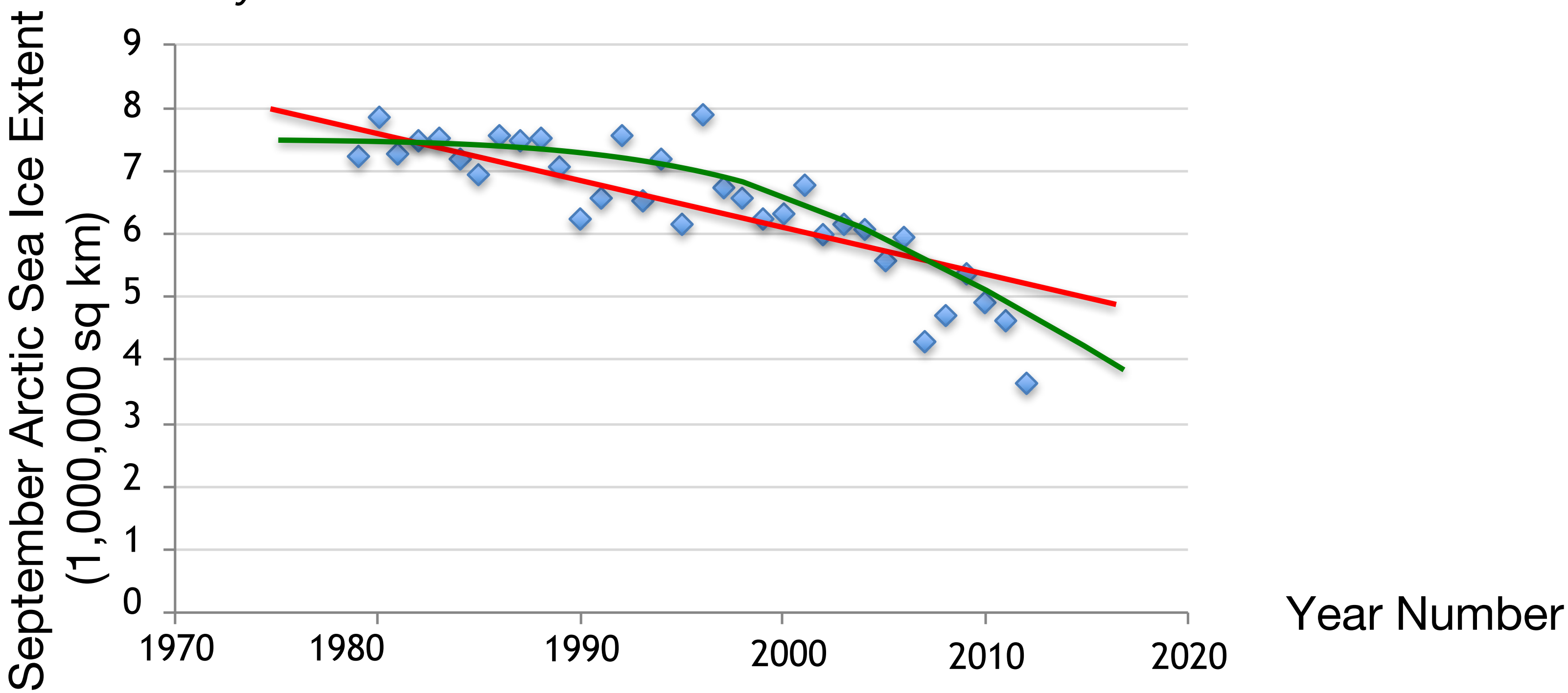


Slide credit: D. Hoiem and L. Lazebnik

Supervised Learning: Regression

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

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

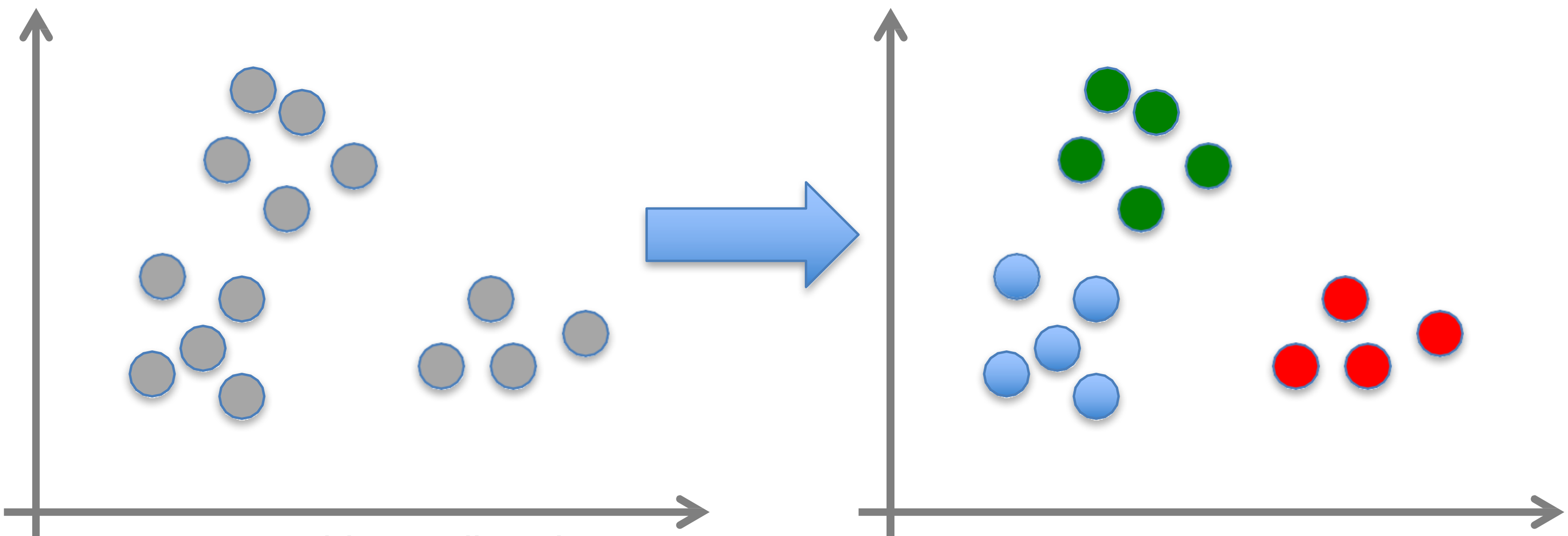


Data from G. Witt. Journal of Statistics Education, Volume 21, Number 1 (2013) Slide Credit: Eric Eaton

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

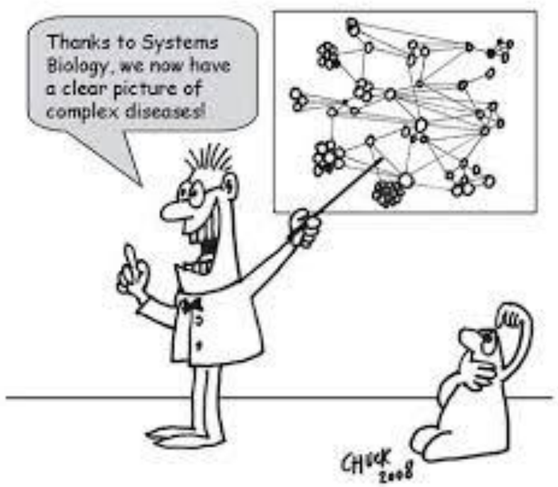


Slide Credit: Eric Eaton

Introduction to Machine Learning

Unsupervised Learning Applications

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



Unsupervised Learning Techniques

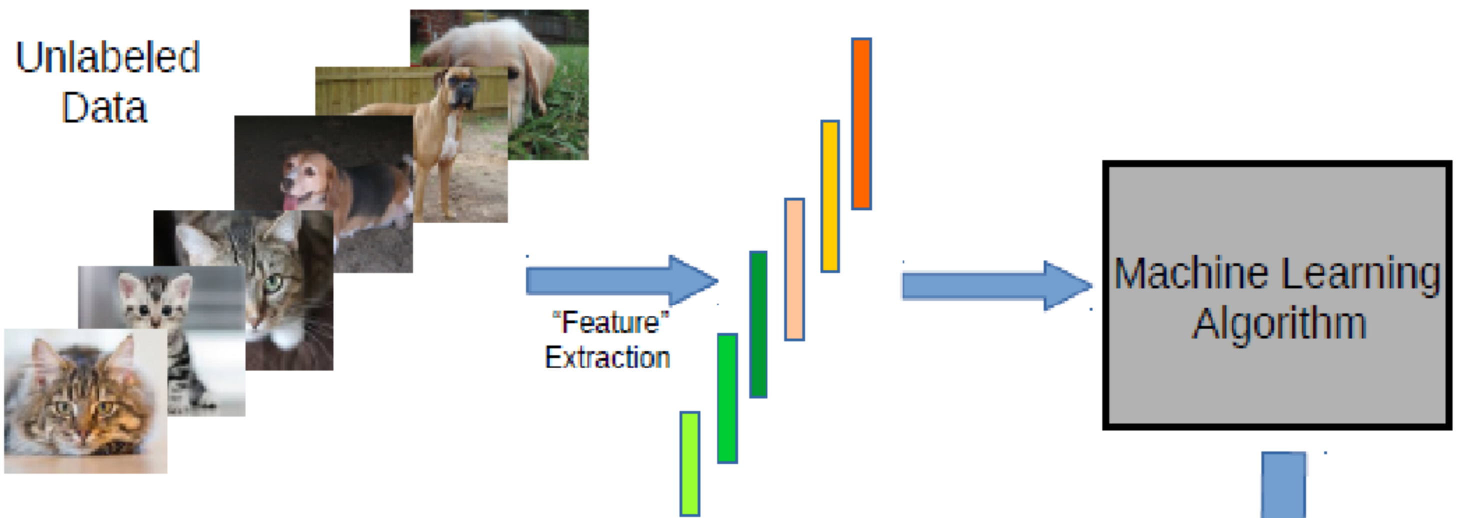
Clustering

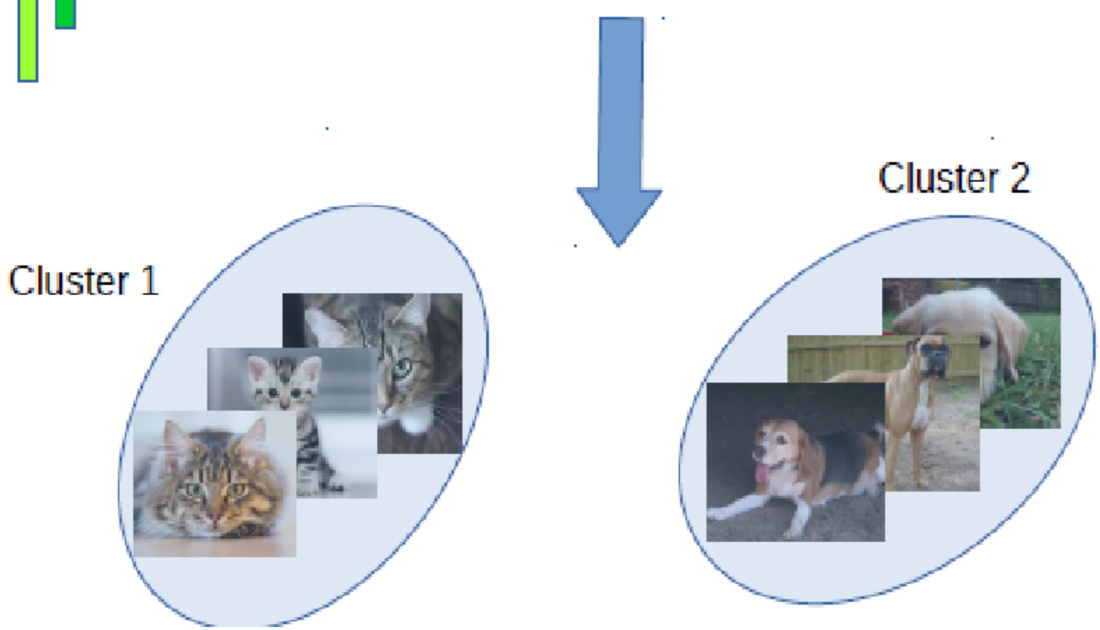
- k-Means
- Hierarchical Cluster Analysis
- Expectation Maximization

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

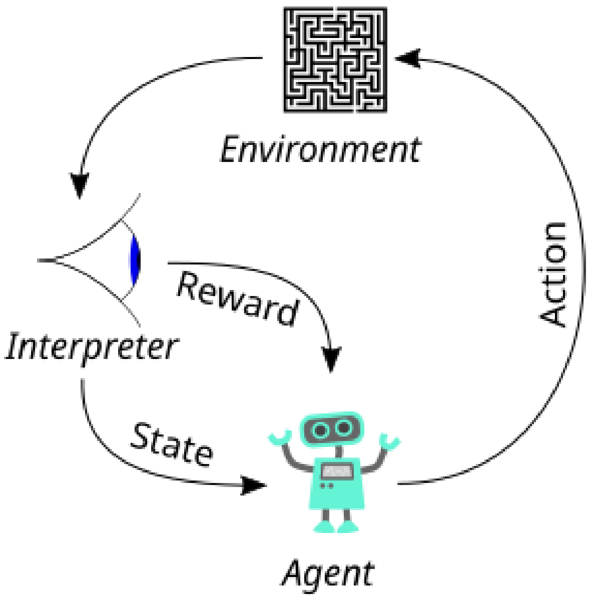
Slide Credit: Piyush Rai

Reinforcement Learning

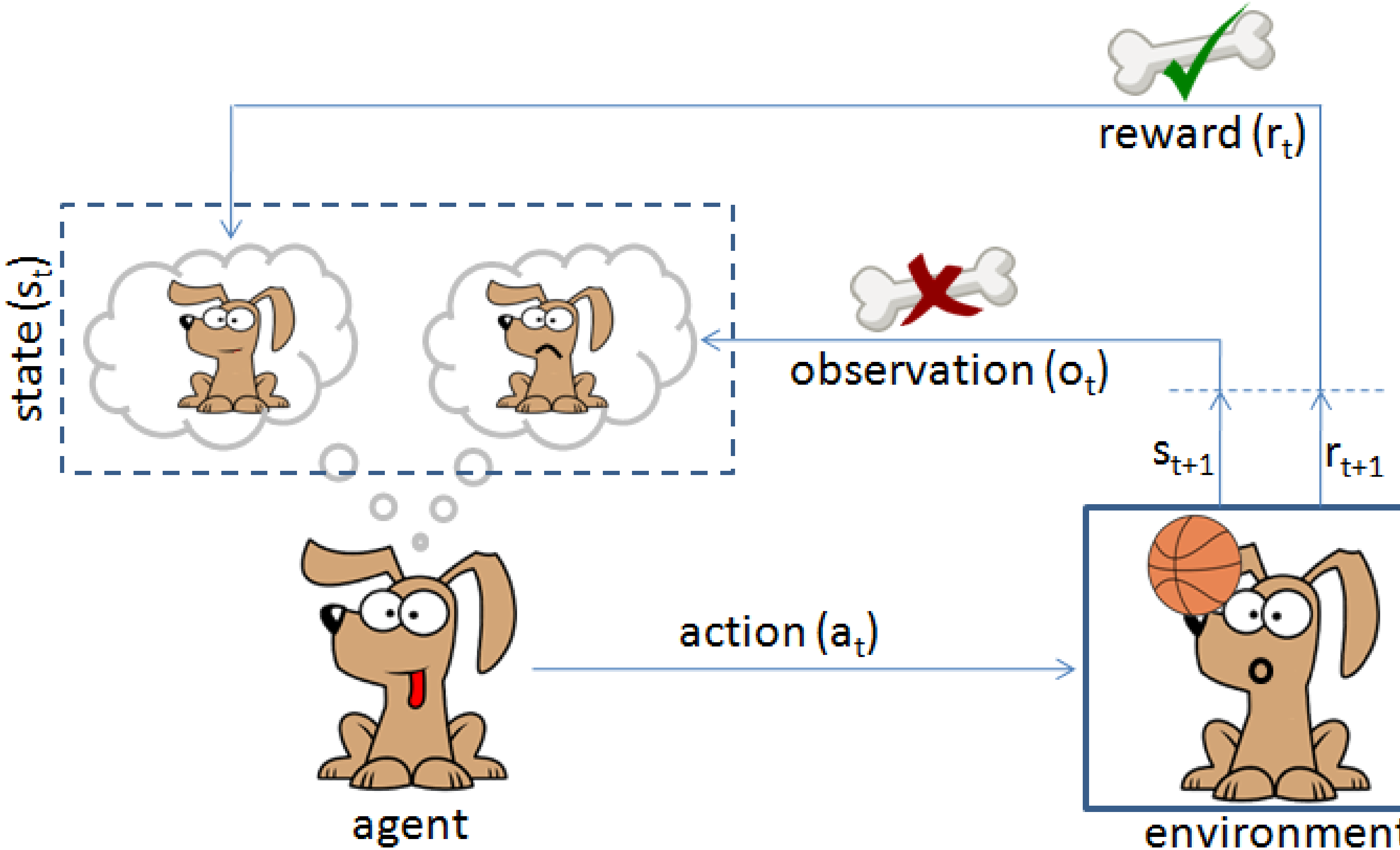
- feedback-based Machine learning technique
- agent learns automatically using feedbacks without any labeled data, unlike supervised learning
- an agent learns to behave in an environment by performing the actions and seeing the results of actions
- For each good action, the agent gets positive feedback, and for each bad action, the agent gets negative feedback or penalty
- Example - How a Robotic dog learns movement of his arms
- solves specific type of problem where decision making is sequential, and the goal is long-term, such as game-playing, robotics, etc.

Reinforcement Learning

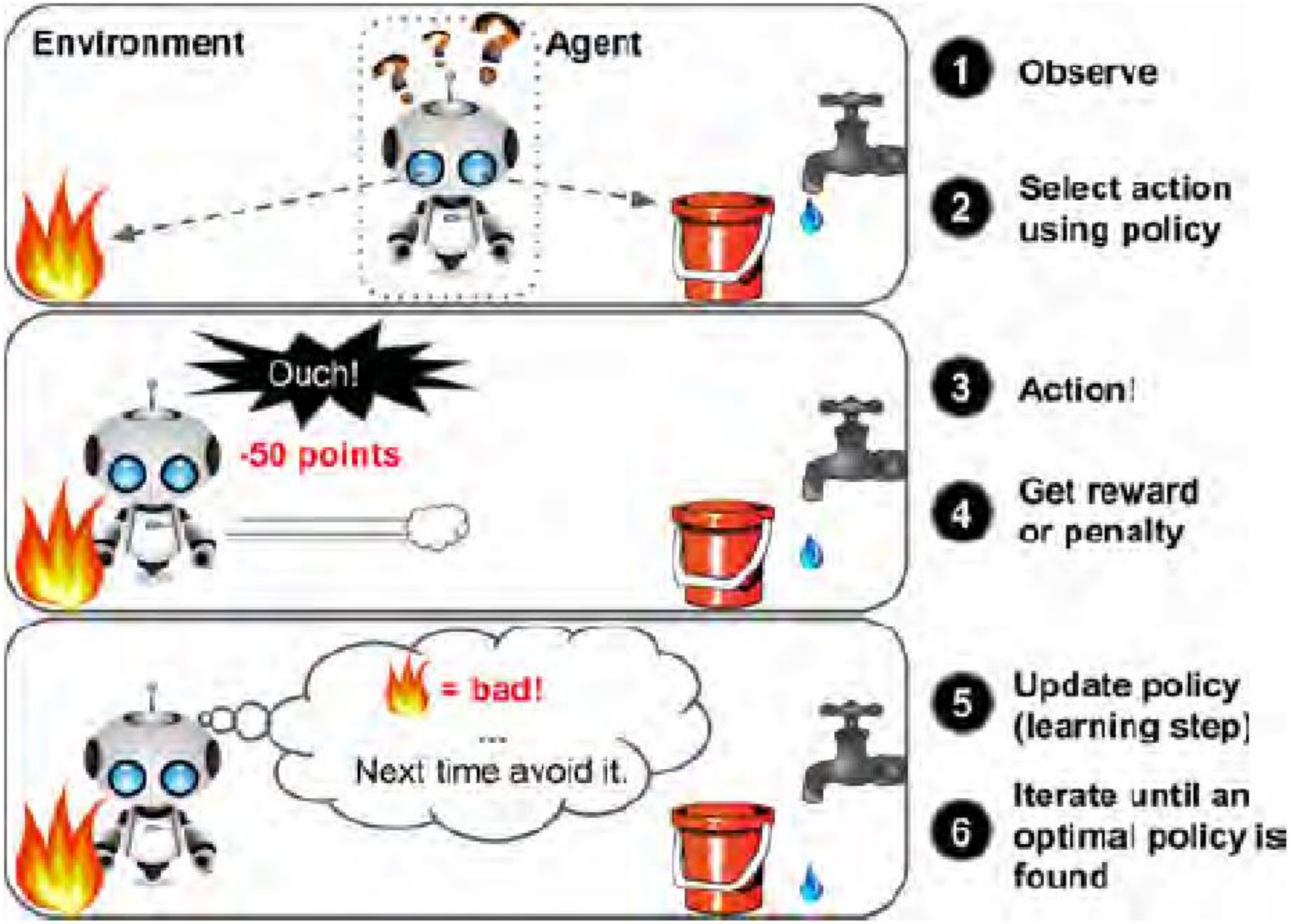
- RL cantered around a digital agent who is put in a specific environment to learn
- Similar to way that we learn new things, agent faces a game-like situation
- must make a series of decisions to try to achieve correct outcome
- Through trial and error, agent will learn what to do (and what not to do) and is rewarded and punished accordingly
- Every time it receives a reward, it reinforces the behaviour and signals the agent to employ the same tactics again next time



Reinforcement Learning



Reinforcement Learning



Summary : Types of Learning

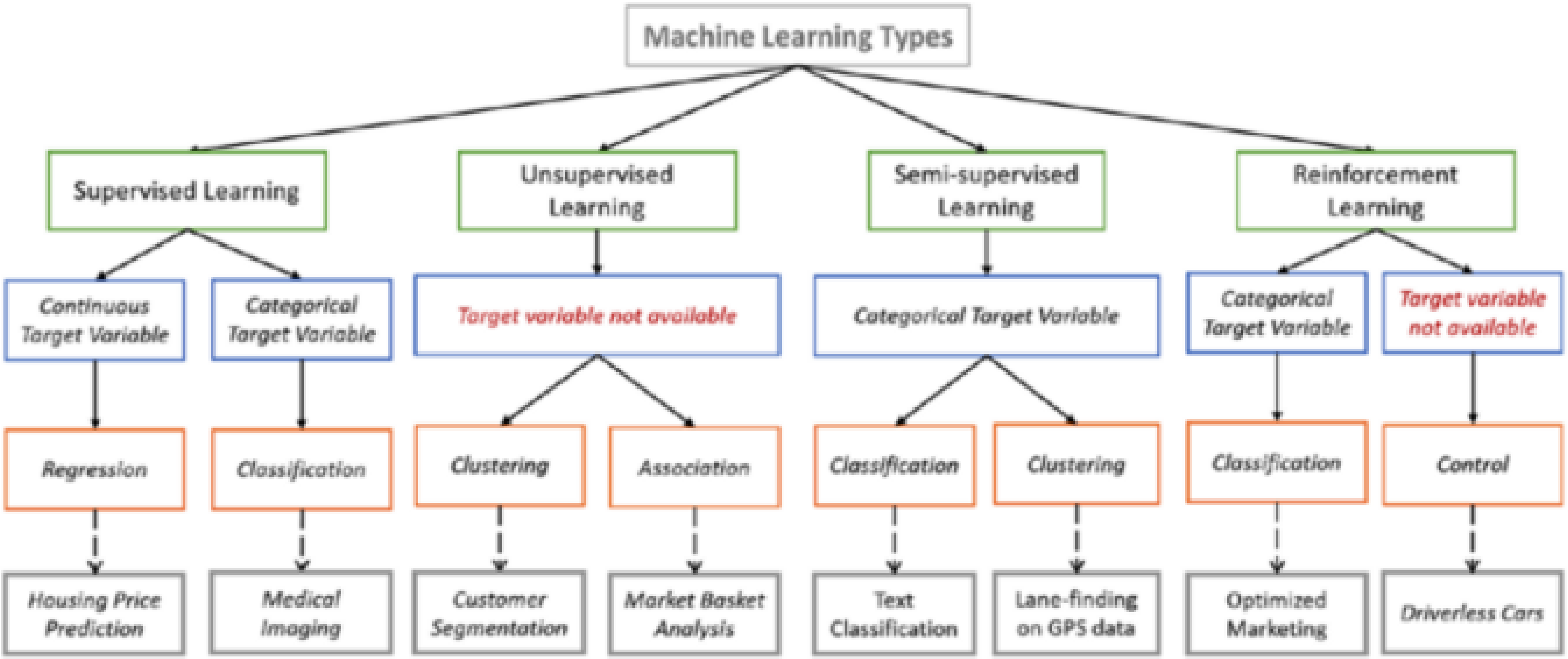


Image courtesy: https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861

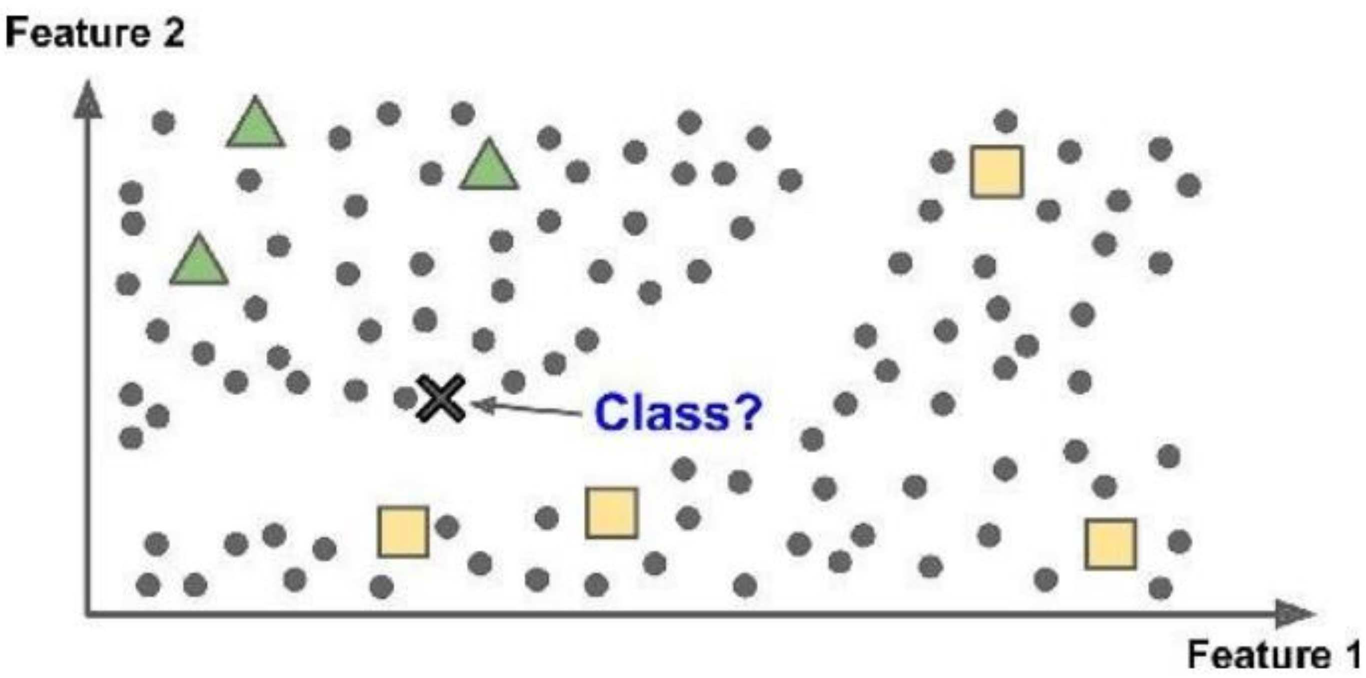
Comparison : Supervised, Unsupervised and Reinforcement Learning

Criteria	Supervised Learning	Unsupervised Learning	Reinforcement Learning
Definition	The machine learns by using labeled data	The machine is trained on unlabeled data without any guidance	An agent interacts with its environment by performing actions & learning from errors or rewards
Type of problems	Regression & classification	Association & clustering	Reward-based
Type of data	Labeled data	Unlabeled data	No predefined data
Training	External supervision	No supervision	No supervision
Approach	Maps the labeled inputs to the known outputs	Understands patterns & discovers the output	Follows the trial-and-error method

Semi supervised Learning

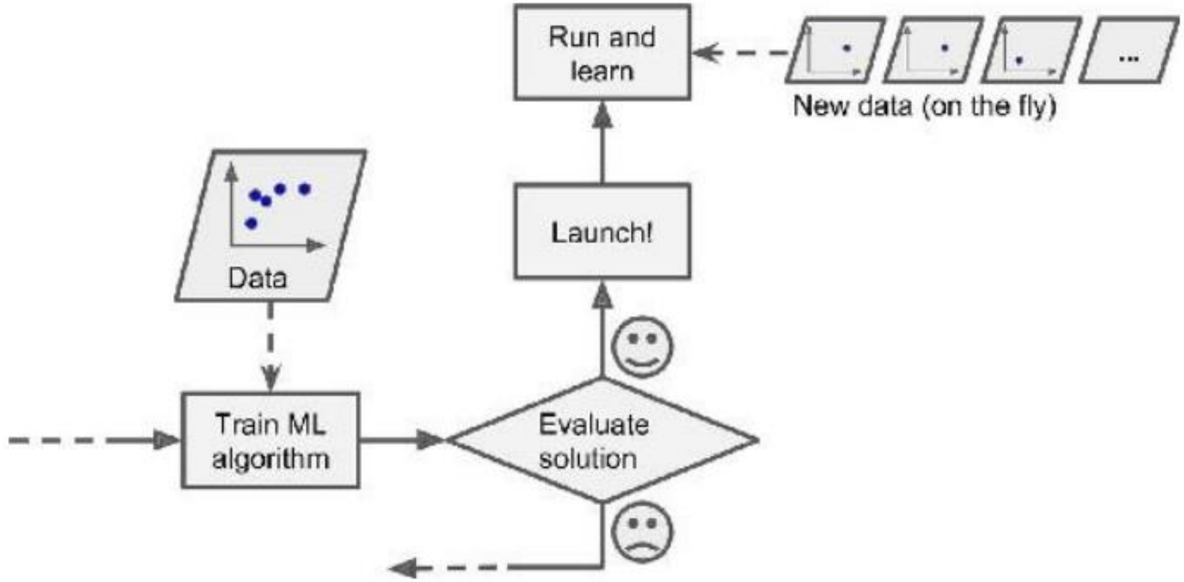
Partially labelled data – some labelled data and a lot of unlabelled data

- Combines unsupervised and supervised learning algorithms
- Photo hosting service, e.g., google photos



Types: Based on how training data is used

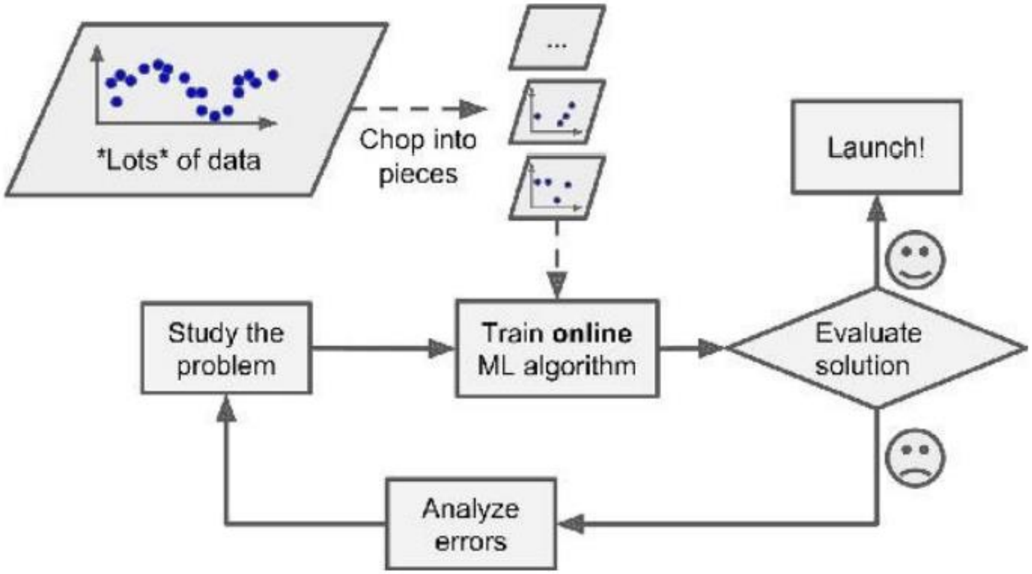
- Batch learning: Uses all available data at a time during training
- Mini Batch learning: Uses a subset of available at a time during training
- Online (incremental) learning: Uses single training data instance at a time during training



Introduction to Machine Learning

Types: Based on how training data is used

- Batch learning: Uses all available data at a time during training
- Mini Batch learning: Uses a subset of available at a time during training
- Online (incremental) learning: Uses single training data instance at a time during training

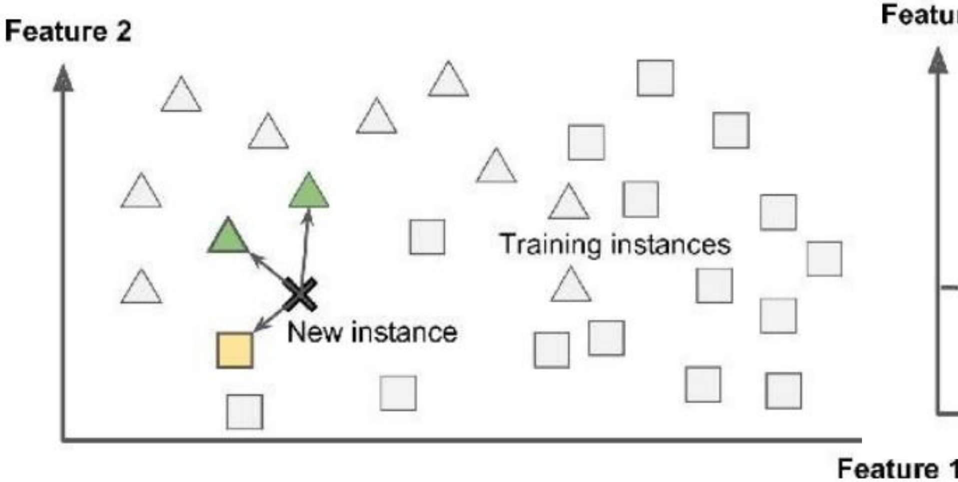


Introduction to Machine Learning

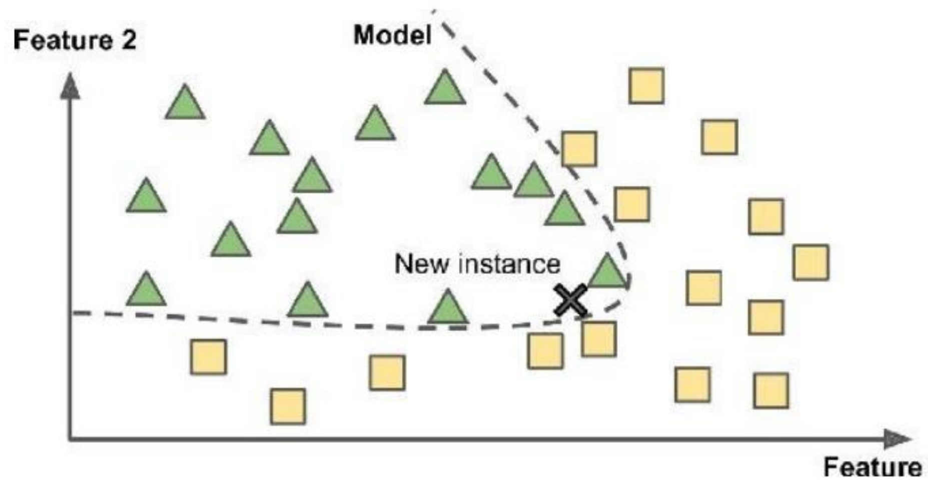
Types: Based on how training data is used

- Instance Based Learning: Compare new data points to known data points
- Model Based learning : Detect patterns in the training data and build a predictive model

Instance Based



Model Based



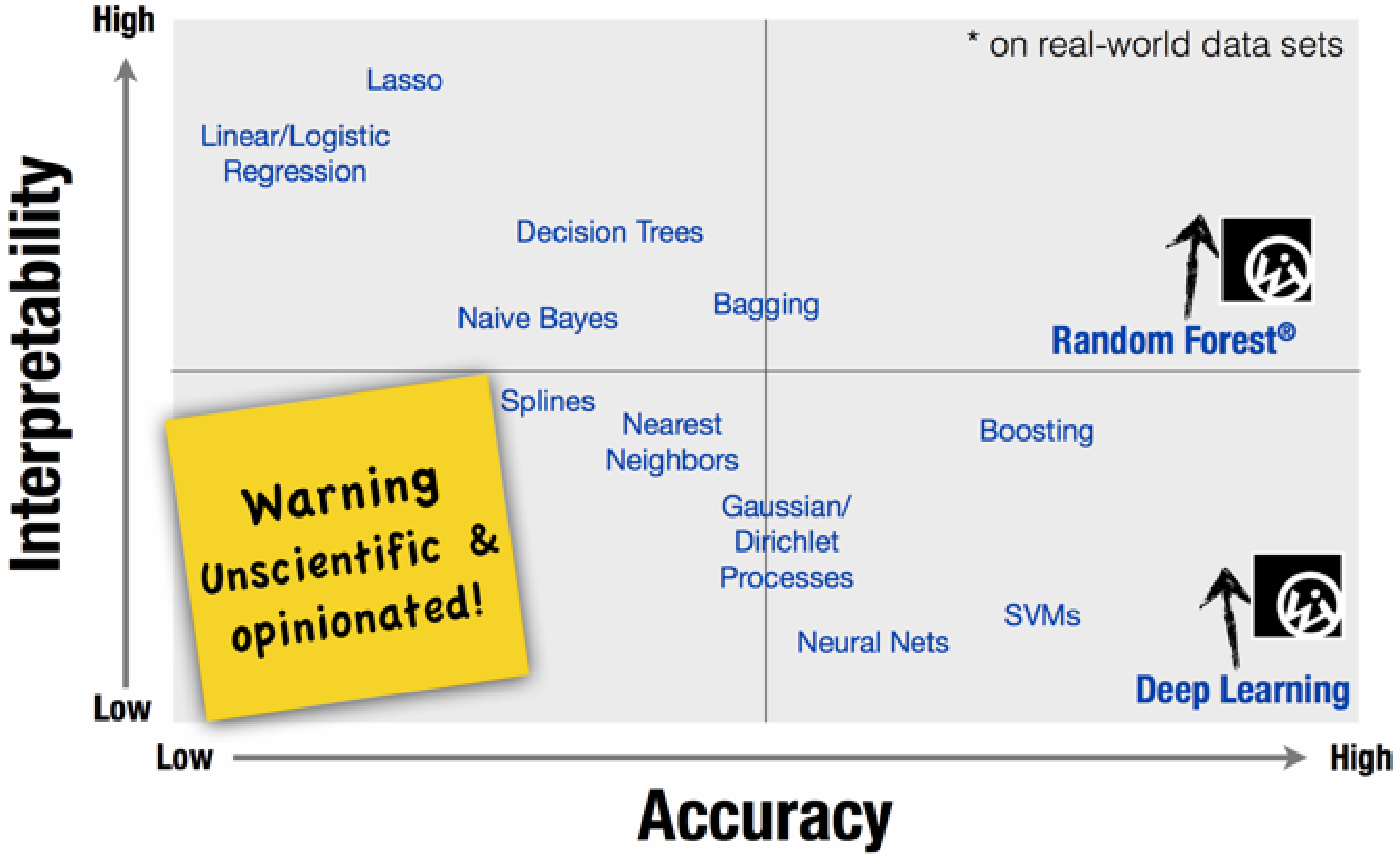
Open source ML programming tools

	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

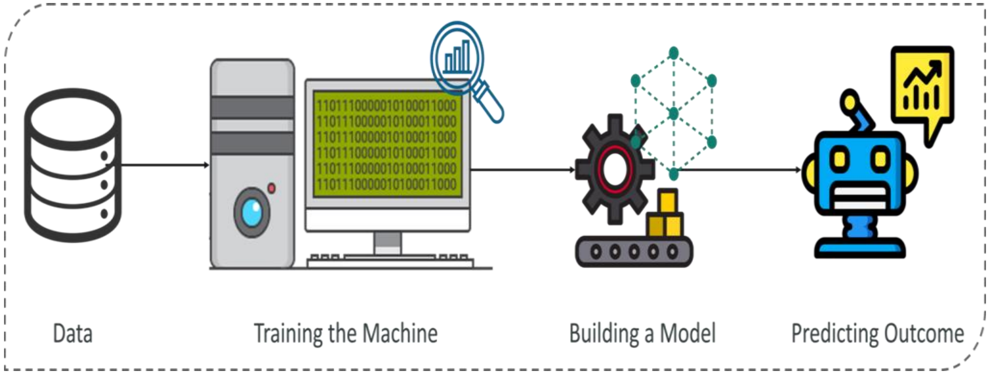
Open source ML programming tools

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

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

- 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

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 !