

Machine Learning Algorithms

Tushar B. Kute,
<http://tusharkute.com>



What is Data Science ?

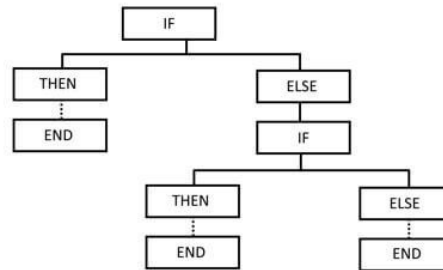
- Data science is an **interdisciplinary** field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from structured and unstructured data, and apply **knowledge** and **actionable insights** from data across a broad range of application domains.
- Data science is related to data **mining**, **machine learning** and **big data**.
- Data science (DS) is a **multidisciplinary** field of study with goal to address the challenges in big data.
- Data science **principles** apply to all data – big and small.

Artificial Intelligence

- Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions.
- The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving.

Major AI Approaches

- Two Major AI Techniques
 - Logic and Rules-Based Approach



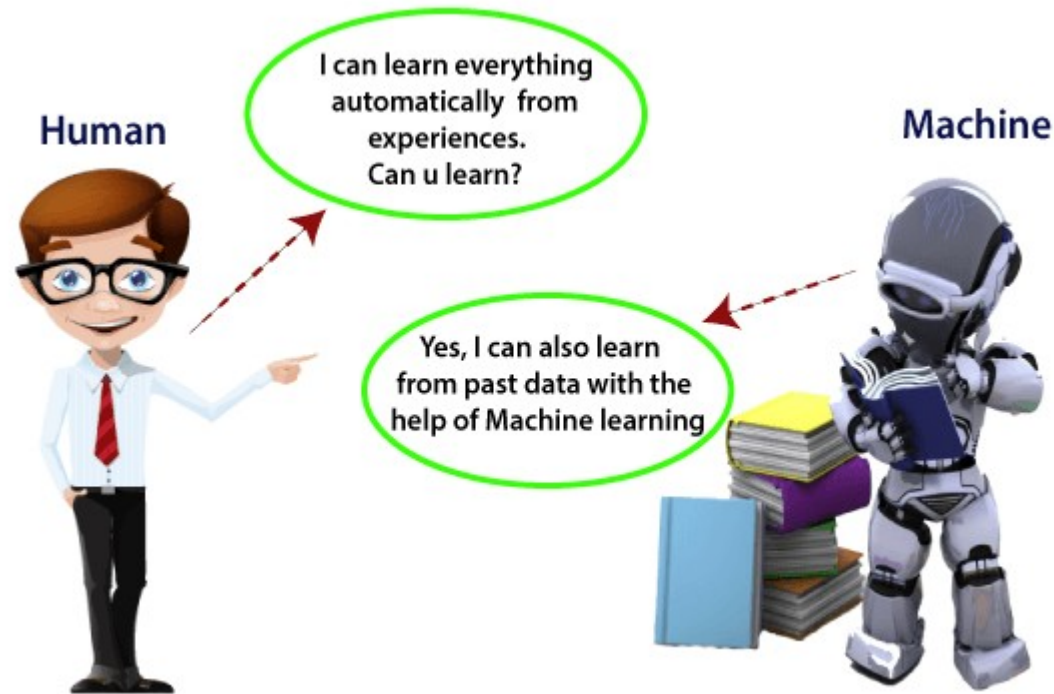
- Machine Learning (Pattern-Based Approach)



Machine Learning

- Machine learning is an application of **artificial intelligence** (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.
- The process of learning begins with **observations** or data, such as examples, **direct experience**, or **instruction**, in order to look for patterns in data and make better decisions in the future based on the examples that we provide.
- The primary aim is to allow the computers learn automatically **without** human intervention or assistance and adjust actions accordingly.

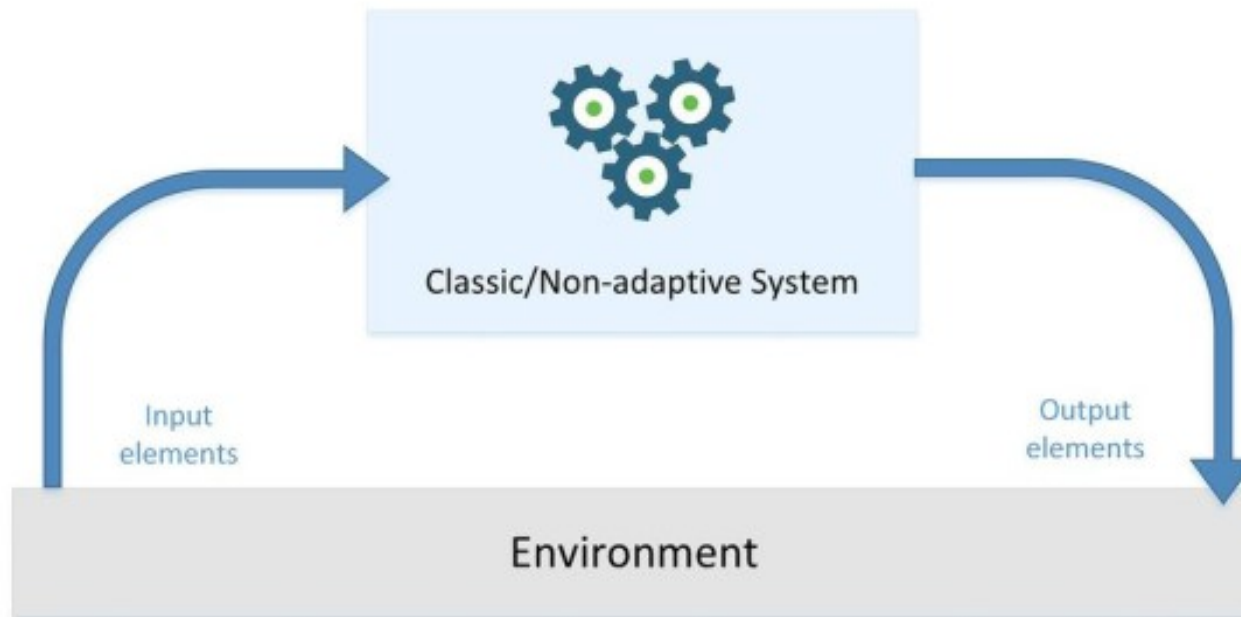
Machine Learning



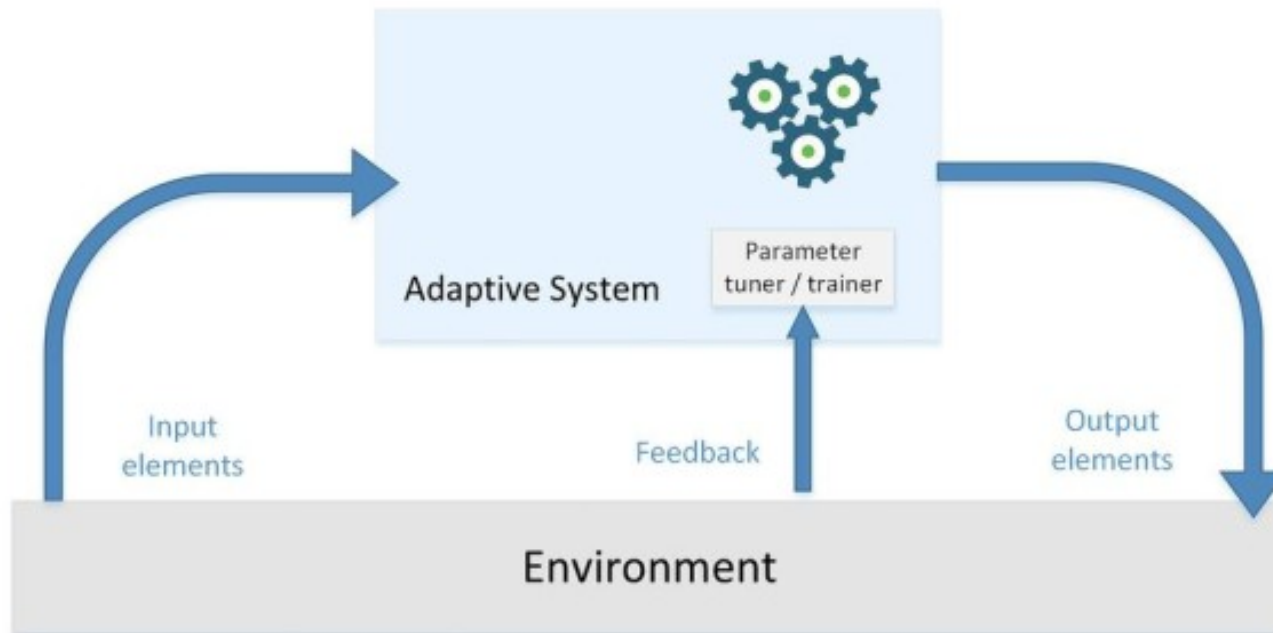
Origins of Machine Learning

- The earliest databases recorded information from the observable environment.
- Astronomers recorded patterns of planets and stars; biologists noted results from experiments crossbreeding plants and animals; and cities recorded tax payments, disease outbreaks, and populations.
- Each of these required a human being to first observe and second, record the observation.
- Today, such observations are increasingly automated and recorded systematically in ever-growing computerized databases.

Classic Systems



Adaptive Systems



Machine Learning

Traditional Programming

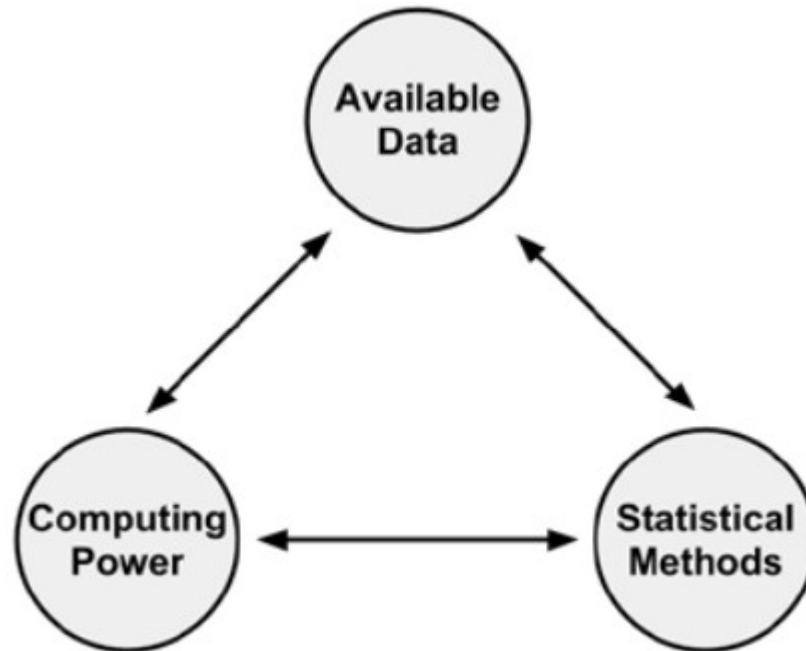


Machine Learning

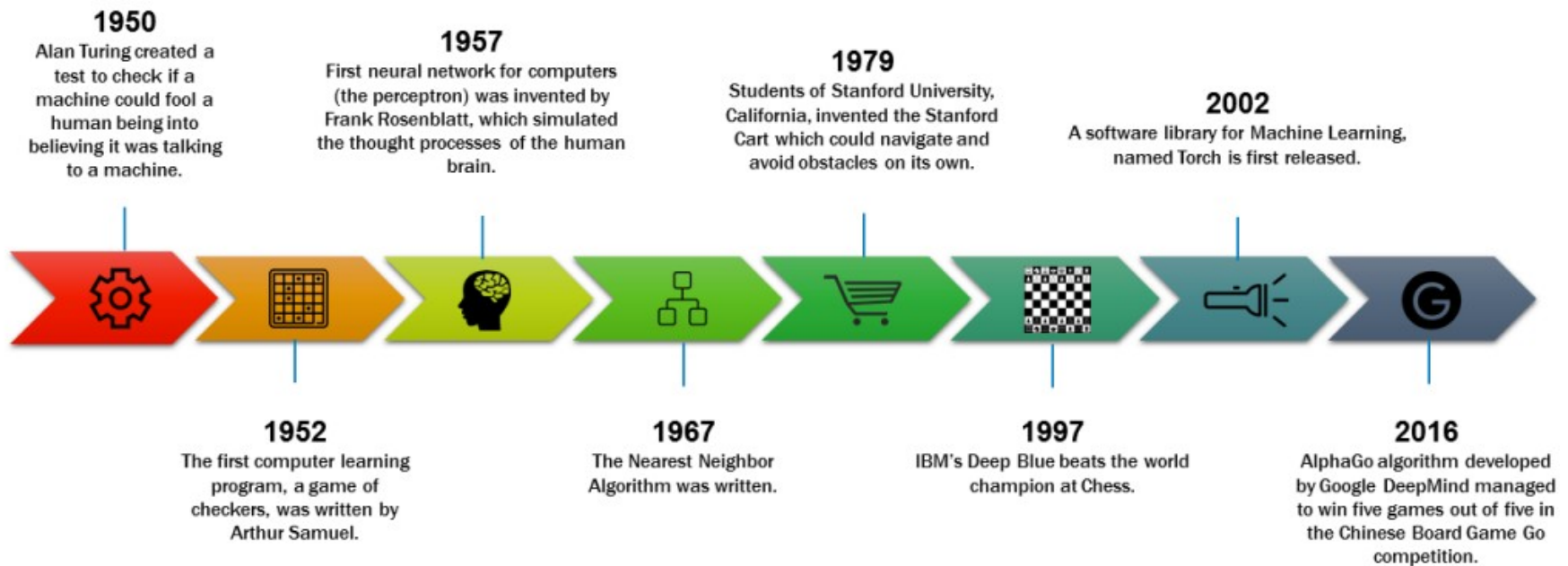


Machine Learning

- The field of study interested in the development of computer algorithms for transforming data into intelligent action is known as machine learning.



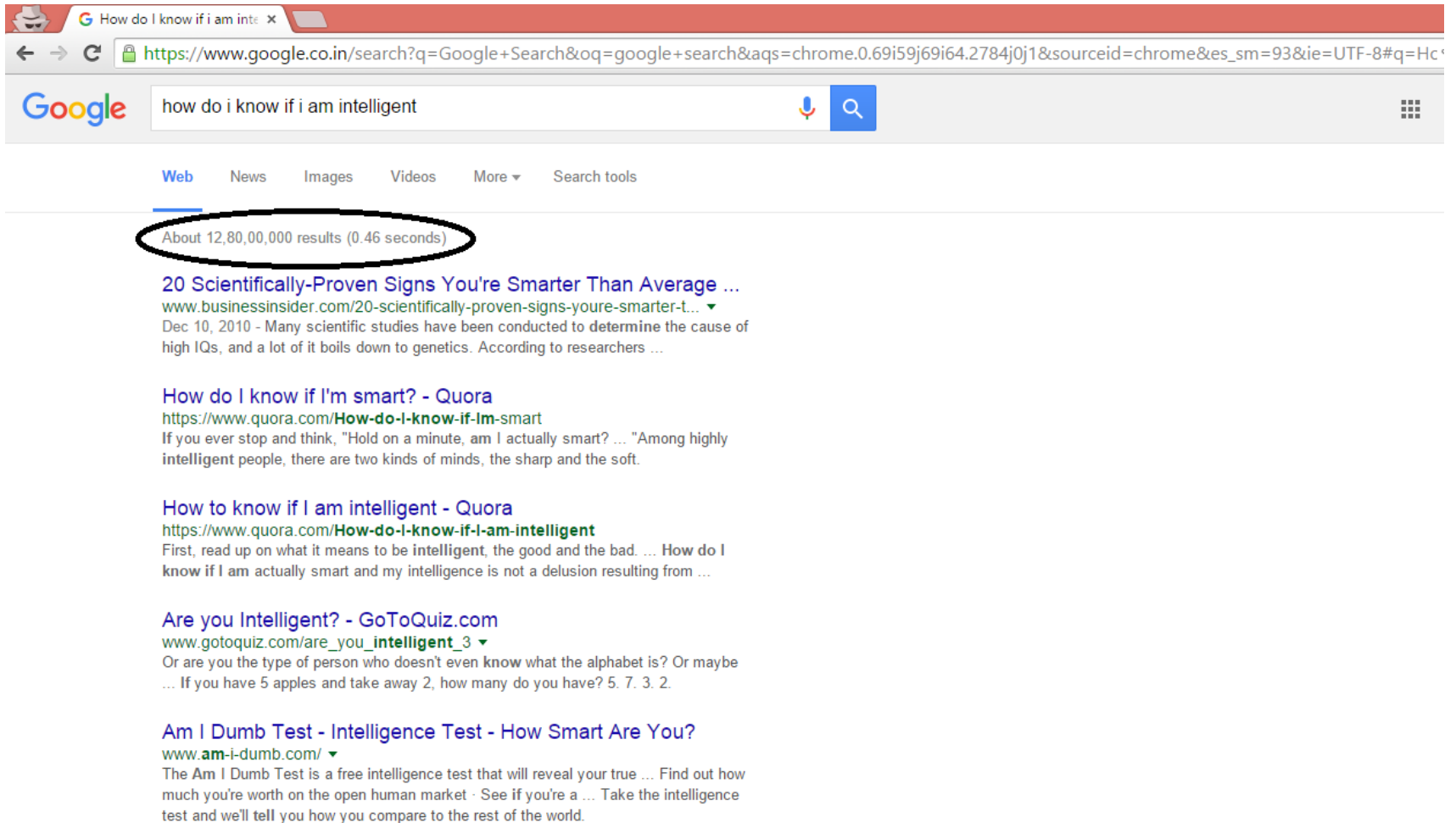
Timeline



Real Life Examples

- Internet Search
- Digital Advertisements (Targeted Advertising and re-targeting)
- Recommender Systems
- Image Recognition
- Speech Recognition
- Gaming
- Price Comparison Websites
- Airline Route Planning
- Fraud and Risk Detection
- Delivery logistics

Internet Search



How do I know if i am inte x

← → ↻ https://www.google.co.in/search?q=Google+Search&oq=google+search&aqs=chrome.0.69i59j69i64.2784j0j1&sourceid=chrome&es_sm=93&ie=UTF-8#q=Hc

Google how do i know if i am intelligent

Web News Images Videos More Search tools

About 12,80,00,000 results (0.46 seconds)

20 Scientifically-Proven Signs You're Smarter Than Average ...
www.businessinsider.com/20-scientifically-proven-signs-youre-smarter-t...
Dec 10, 2010 - Many scientific studies have been conducted to determine the cause of high IQs, and a lot of it boils down to genetics. According to researchers ...

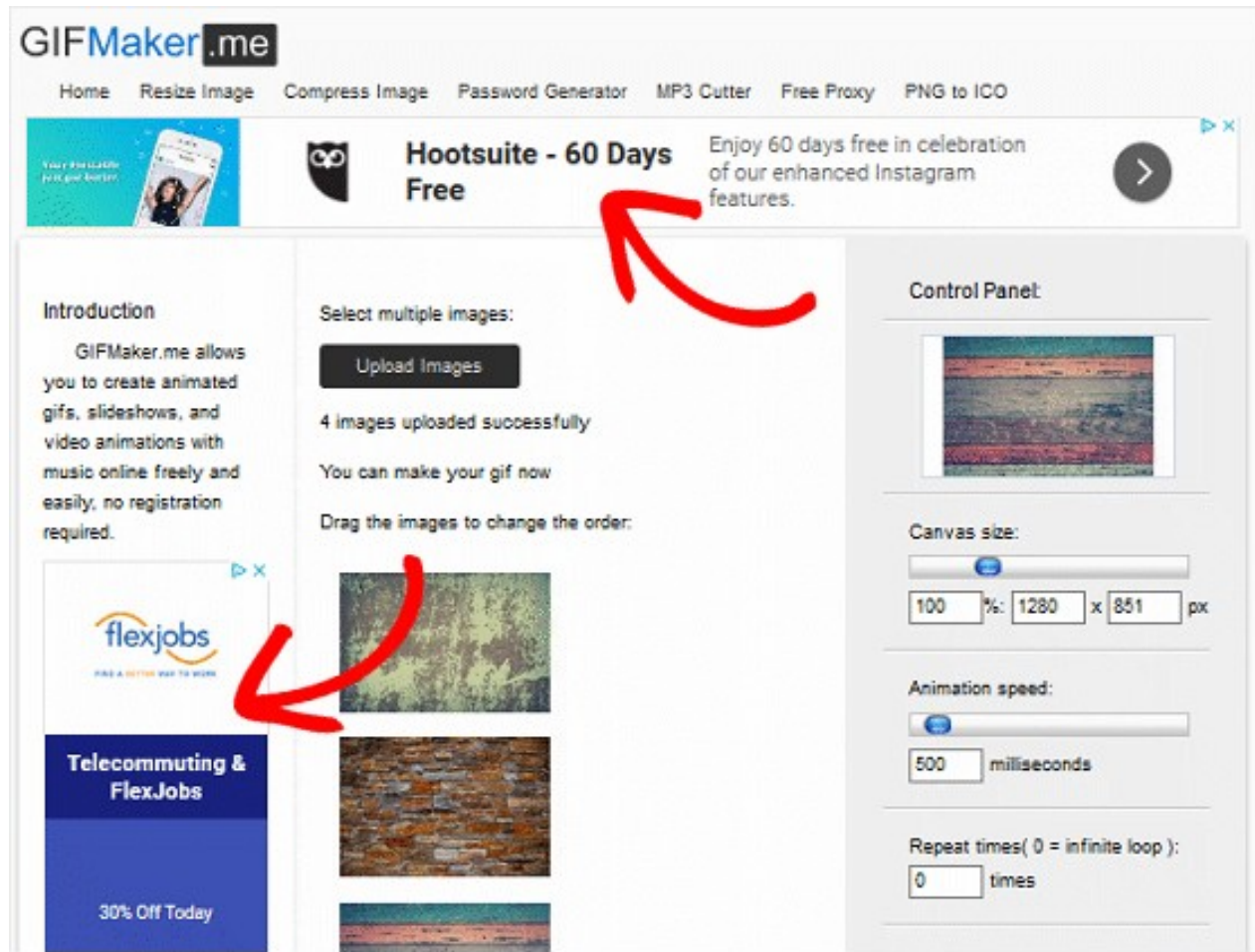
How do I know if I'm smart? - Quora
<https://www.quora.com/How-do-I-know-if-Im-smart>
If you ever stop and think, "Hold on a minute, am I actually smart? ... "Among highly intelligent people, there are two kinds of minds, the sharp and the soft.

How to know if I am intelligent - Quora
<https://www.quora.com/How-do-I-know-if-I-am-intelligent>
First, read up on what it means to be intelligent, the good and the bad. ... How do I know if I am actually smart and my intelligence is not a delusion resulting from ...

Are you Intelligent? - GoToQuiz.com
www.gotoquiz.com/are_you_intelligent_3
Or are you the type of person who doesn't even know what the alphabet is? Or maybe ... If you have 5 apples and take away 2, how many do you have? 5. 7. 3. 2.

Am I Dumb Test - Intelligence Test - How Smart Are You?
www.am-i-dumb.com/
The Am I Dumb Test is a free intelligence test that will reveal your true ... Find out how much you're worth on the open human market - See if you're a ... Take the intelligence test and we'll tell you how you compare to the rest of the world.

Targeting Advertisement



GIFMaker.me

Home Resize Image Compress Image Password Generator MP3 Cutter Free Proxy PNG to ICO

Hootsuite - 60 Days Free Enjoy 60 days free in celebration of our enhanced Instagram features.

Introduction

GIFMaker.me allows you to create animated gifs, slideshows, and video animations with music online freely and easily, no registration required.

Select multiple images:
Upload Images

4 images uploaded successfully

You can make your gif now

Drag the images to change the order:

flexjobs
FIND A BETTER WAY TO WORK

Telecommuting & FlexJobs

30% Off Today

Control Panel:

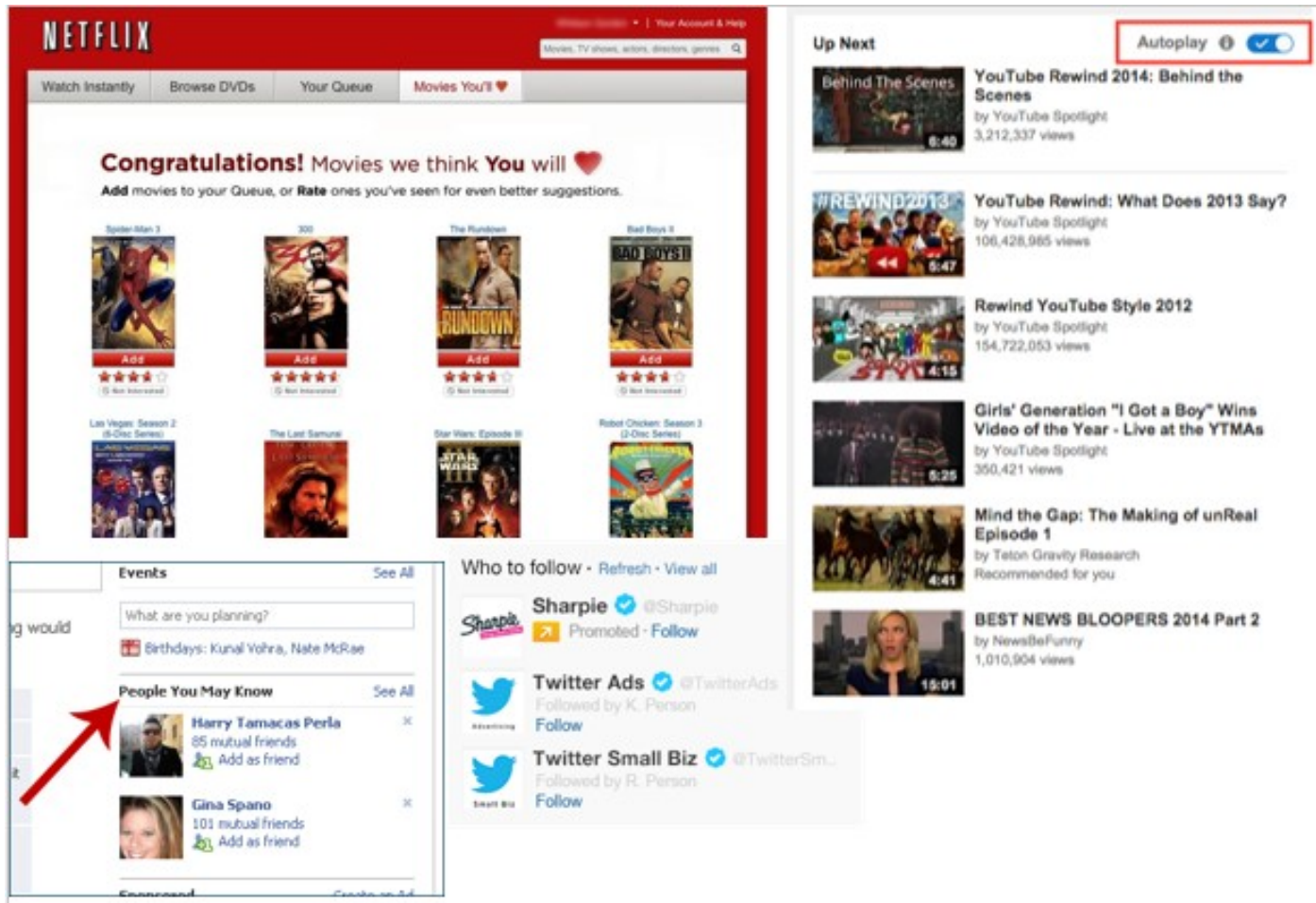
Canvas size: 100 %: 1280 x 851 px

Animation speed: 500 milliseconds

Repeat times (0 = infinite loop): 0 times

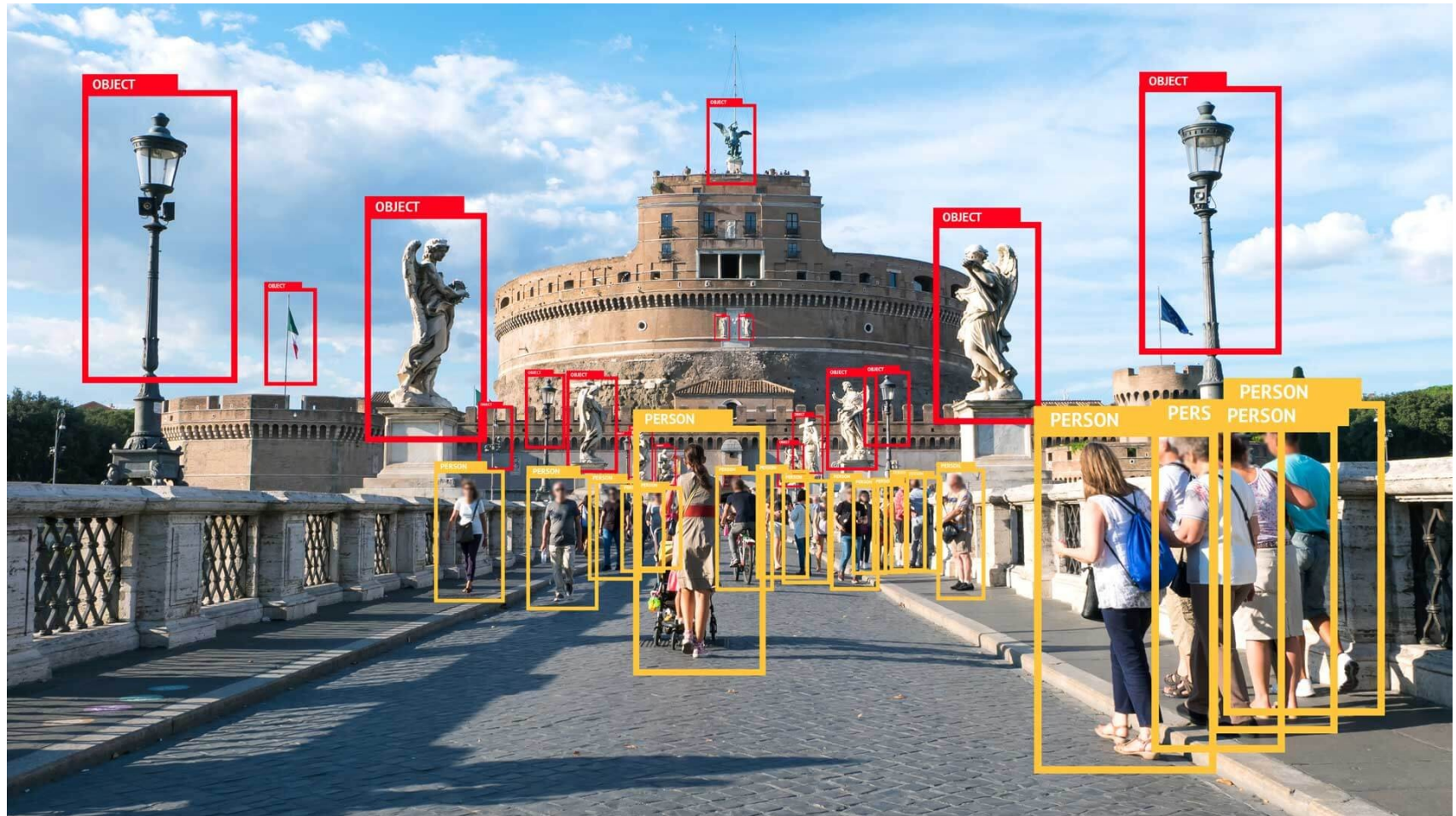
Two red arrows point to the Hootsuite advertisement and the flexjobs advertisement.

Recommender System



The image displays two screenshots illustrating recommender systems. The left screenshot shows the Netflix interface with a red header and navigation tabs. A central section titled "Congratulations! Movies we think You will" features a grid of movie thumbnails with "Add" buttons and star ratings. Below this, a sidebar shows "Events" and "People You May Know" with a red arrow pointing to the latter. The right screenshot shows a YouTube "Up Next" playlist with a red box around the "Autoplay" toggle, which is turned on. The playlist includes videos like "Behind The Scenes" and "YouTube Rewind 2014: Behind the Scenes".

Image Recognition



Speech Recognition



Computer Games



Price Comparison Website



trivago

Hotels in New York

Website A
25 Hotels

Website B
56 Hotels

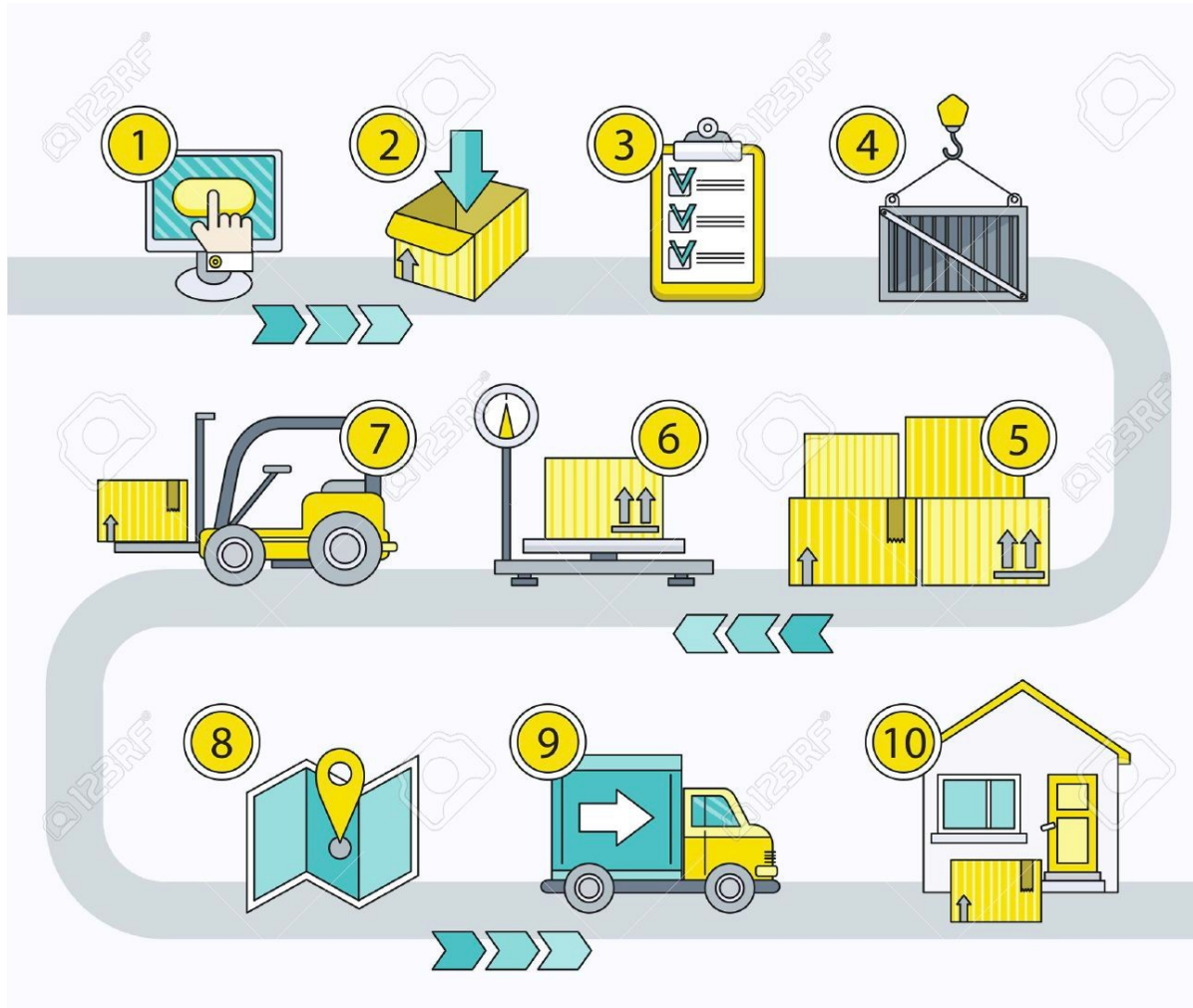
Airline Route Planning



Fraud Detection



Delivery Logistics



Uses

- Predict the outcomes of elections
- Identify and filter spam messages from e-mail
- Foresee criminal activity
- Automate traffic signals according to road conditions
- Produce financial estimates of storms and natural disasters
- Examine customer churn
- Create auto-piloting planes and auto-driving cars
- Stock market prediction
- Target advertising to specific types of consumers

सकाळ

विद्यापीठात विद्यार्थ्यांचा 'एक्झिट पोल' 'रँडम फॉरेस्ट मॉडेल'नुसार युतीच राज्यात आघाडीवर

पुणे, ता. २१ : राज्यात भाजप आणि शिवसेना युती आघाडीवर असेल, असा अंदाज वर्तविणाऱ्या चाचण्यांचे कल (एक्झिट पोल) नुकतेच प्रसिद्ध झाले आहेत. सावित्रीबाई फुले पुणे विद्यापीठातील विद्यार्थ्यांनीही त्याला दुजोरा दिला आहे. भारतीय जनता पक्षाला १७ ते २३ आणि शिवसेनेला १६ ते २१ जागा मिळतील, असा अंदाज विद्यार्थ्यांनी 'रँडम फॉरेस्ट मॉडेल' पद्धत वापरून वर्तविला आहे. राष्ट्रवादी काँग्रेसला ३ ते ९ व काँग्रेसला १ ते ६ जागा मिळतील, असा अंदाज त्यांनी वर्तवला आहे.

विद्यापीठाच्या संख्याशास्त्र विभागातील एमएस्सी (द्वितीय वर्ष)



करणारे विनय तिवारी, आर. विश्वनाथ, शरद कोळसे या विद्यार्थ्यांनी सहायक प्राध्यापक डॉ. आकांक्षा काशीकर यांच्या मार्गदर्शनाखाली हा अंदाज दिला आहे.

निवडणूक आयोगाच्या संकेतस्थळावरून सर्वेक्षणासाठी लागणारी माहिती त्यांनी मिळविली. जनमानसाचा कल ओळखण्यासाठी 'सीएसडीएस-लोकनीती' सर्वेक्षण अहवालातून नोंदी घेतल्या.

त्याचबरोबर सध्याच्या सरकारच्या कामगिरीबद्दल लोकांच्या प्रतिक्रिया, पंतप्रधानपदाच्या संभाव्य उमेदवारांची लोकप्रियता, मागील निवडणुकीतील आपले मत यंदा बदलू इच्छिणारे मतदार यांचा अभ्यास करण्यात आला. या अंदाजासाठी रँडम फॉरेस्ट मॉडेल वापरण्यापूर्वी २००९ आणि २०१४च्या निवडणुकांचे अंदाज पडताळून पाहण्यात आले. हे अंदाज प्रत्यक्ष निकालांशी पडताळून पाहिले असता, ते जवळपास ९६ टक्के जुळत असल्याचे निदर्शनास आले. म्हणूनच अभ्यासात माहितीच्या विश्लेषणासाठी या पद्धतीचा वापर करण्यात आला, असे डॉ. काशीकर यांनी सांगितले.



संख्याशास्त्र आणि संगणकशास्त्र याची सांगड घालून आणि

मशिन लर्निंगच्या साह्याने उपलब्ध माहितीचे विश्लेषण केले. संख्याशास्त्रातील अभ्यासाची वेगवेगळी मॉडेल्स वापरून १९७७ पासून ते आतापर्यंतच्या लोकसभा आणि विधानसभा निवडणुकीतील माहितीचा अभ्यास केला. त्यामुळे संख्याशास्त्राचा वापर करून वर्तविलेला अंदाज हा निवडणुकीच्या निकालांच्या जवळ जाणारा असेल. - शरद कोळसे, विद्यार्थी

Recognizing patterns

- Pattern recognition is the automated recognition of patterns and regularities in data. It has applications in
 - statistical data analysis,
 - signal processing,
 - image analysis,
 - information retrieval,
 - bioinformatics,
 - data compression,
 - computer graphics and
 - machine learning.

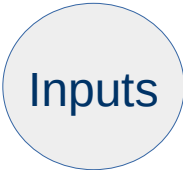
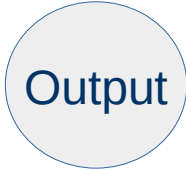
How do machine learn ?

- A commonly cited formal definition of machine learning, proposed by computer scientist Tom M. Mitchell, says that a machine is said to learn if it is able to take experience and utilize it such that its performance improves up on similar experiences in the future.
- This definition is fairly exact, yet says little about how machine learning techniques actually learn to transform data into actionable knowledge.

Training a dataset

- The process of fitting a particular model to a dataset is known as training.
- Why is this not called learning? First, note that the learning process does not end with the step of data abstraction.
- Learning requires an additional step to generalize the knowledge to future data.
- Second, the term training more accurately describes the actual process undertaken when the model is fitted to the data.

Practical Machine Learning

	X	Y	Z	
	5	2	14	
	8	5	22	
	4	8	14	
	9	2	20	
	7	1	15	
	7	8	23	
	Z = ?			

Practical Machine Learning

X	Y	Z	Pre	Error
5	2	14	12	-2
8	5	22	21	-1
4	8	14	16	+2
9	2	20	20	0
7	1	15	15	0
7	8	23	22	-1

$$Z = 2X + Y$$

----> ML Model

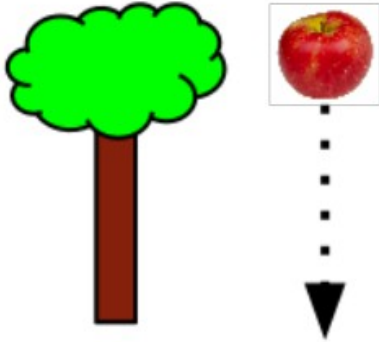
Prediction ----> X = 6 Y = 8 Z = ?

if 20 == 19:

95%

Training a dataset

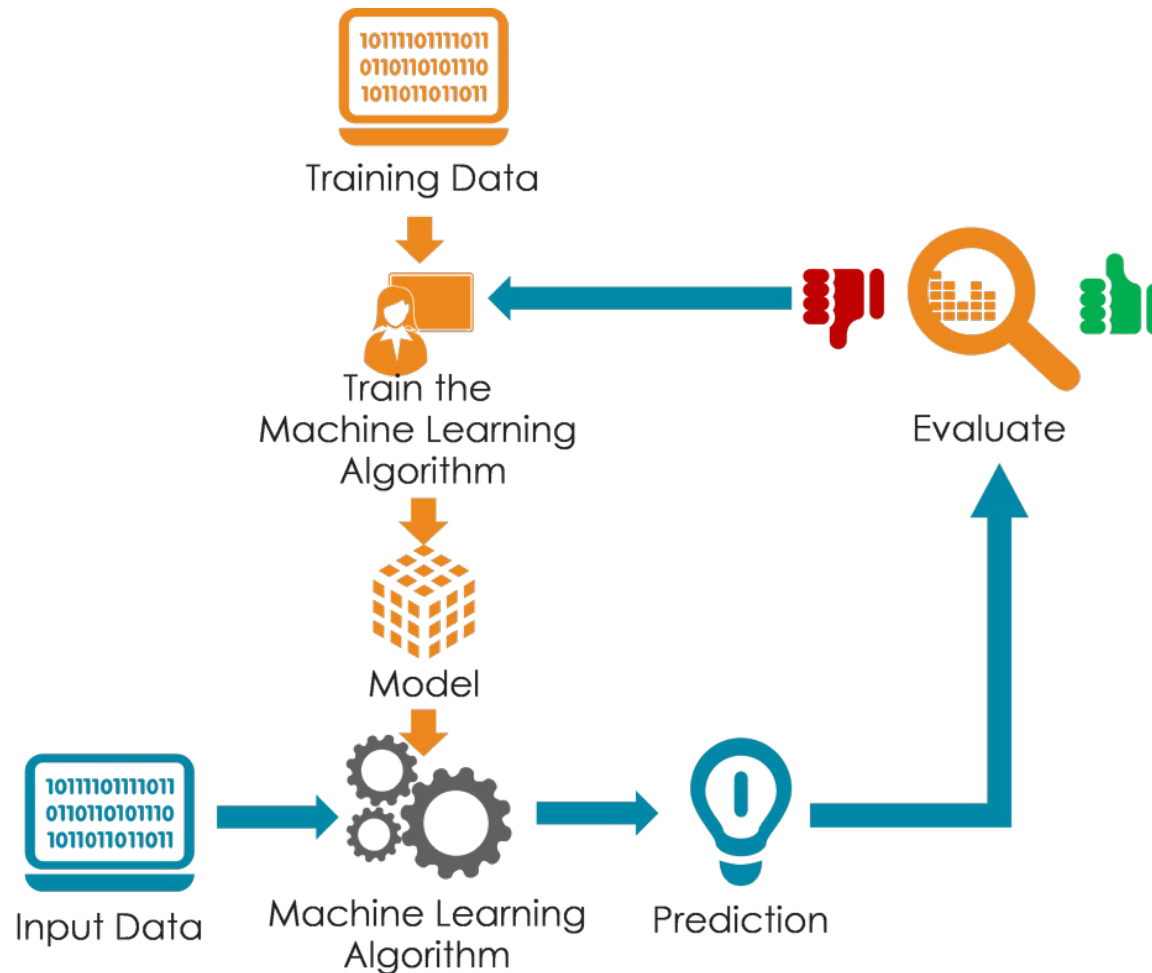
Observations → Data → Model



velocity	time
9.8	1
39.2	2
88.2	3
156.8	4
245	5

$$g = 9.8 \text{ m/s}^2$$

Training a dataset



Well-Posed Learning Problems

- 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 .

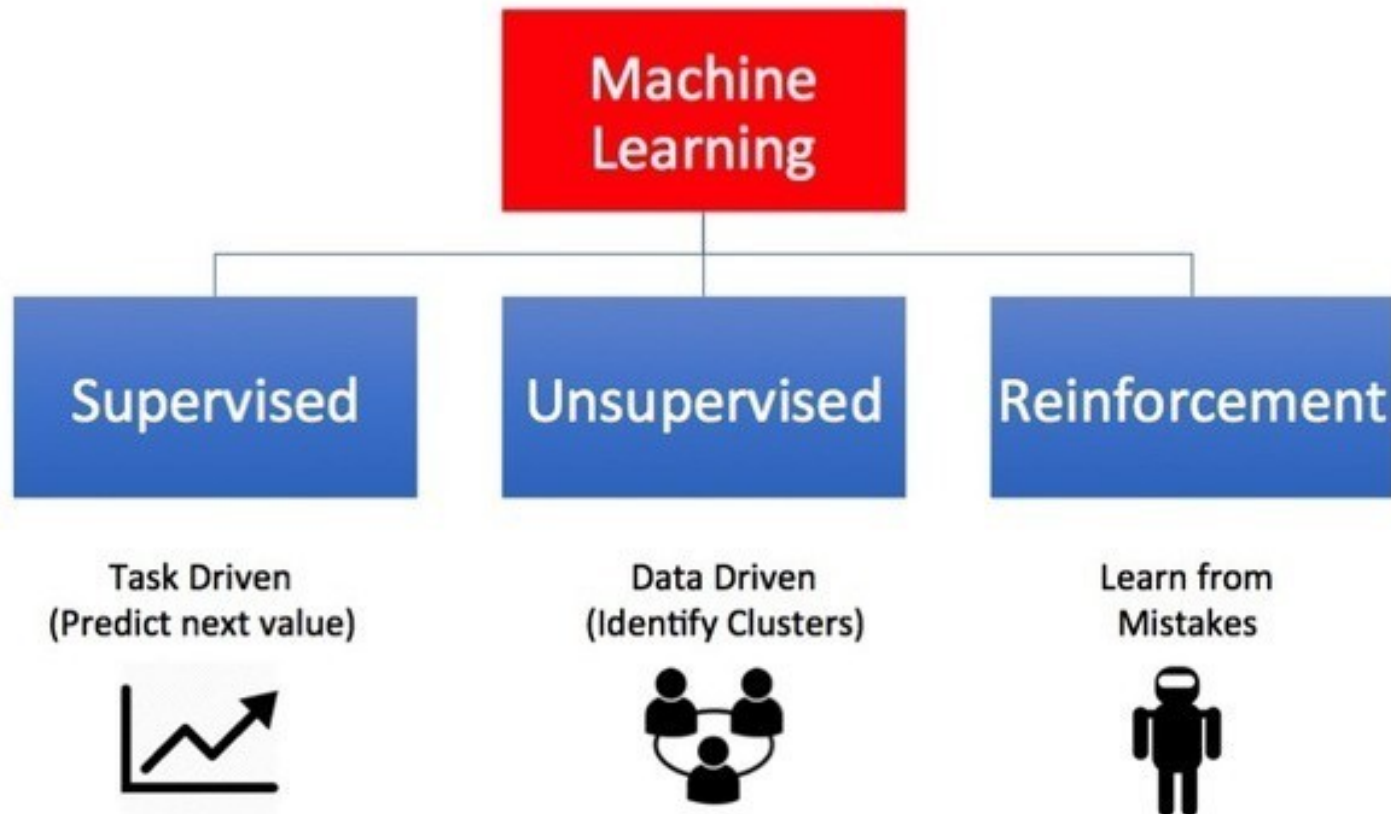
Well-Posed Learning Problems - Examples

- A checkers learning problem
 - Task T : playing checkers
 - Performance measure P : percent of games won against opponents
 - Training experience E : playing practice games against itself
- A handwriting recognition learning problem
 - Task T : recognizing and classifying handwritten words within images
 - Performance measure P : percent of words correctly classified
 - Training experience E : a database of handwritten words with given classifications

Well-Posed Learning Problems - Examples

- A robot driving learning problem
 - Task T : driving on public four-lane highways using vision sensors
 - Performance measure P : average distance traveled before an error (as judged by human overseer)
 - Training experience E : a sequence of images and steering commands recorded while observing a human driver

Types of Machine Learning



Supervised Learning

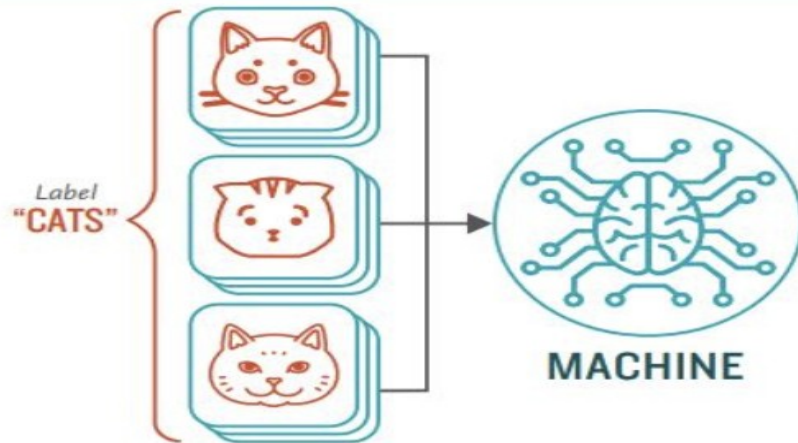
- Supervised learning (SL) is the machine learning task of learning a function that maps an input to an output based on example input-output pairs.
- It infers a function from labeled training data consisting of a set of training examples.
- In supervised learning, each example is a pair consisting of an input object (typically a vector) and a desired output value (also called the supervisory signal).
- A supervised learning algorithm analyzes the training data and produces an inferred function, which can be used for mapping new examples.

Supervised Machine Learning

How **Supervised** Machine Learning Works

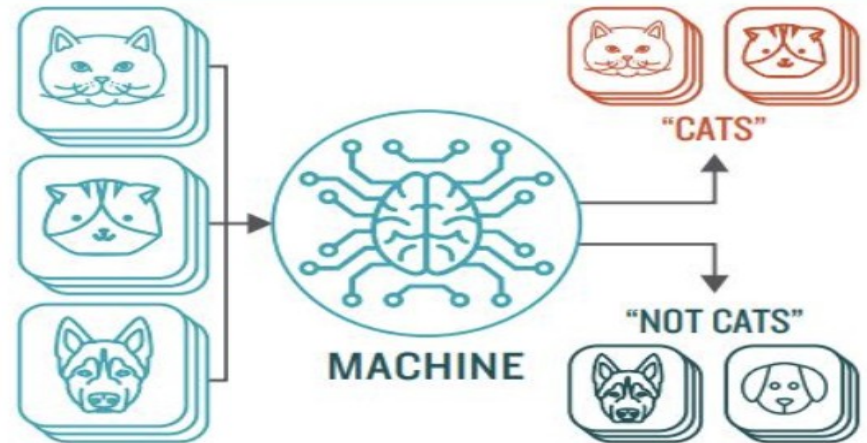
STEP 1

Provide the machine learning algorithm categorized or "labeled" input and output data from to learn

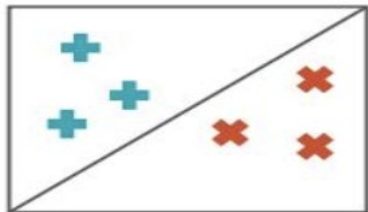


STEP 2

Feed the machine new, unlabeled information to see if it tags new data appropriately. If not, continue refining the algorithm

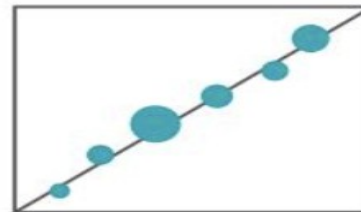


TYPES OF PROBLEMS TO WHICH IT'S SUITED



CLASSIFICATION

Sorting items into categories



REGRESSION

Identifying real values (dollars, weight, etc.)

Supervised Learning : Examples

- Support-vector machines
- Linear regression
- Logistic regression
- Naive Bayes
- Linear discriminant analysis
- Decision trees
- K-nearest neighbor algorithm
- Neural networks (Multilayer Perceptron)

Un-Supervised Learning

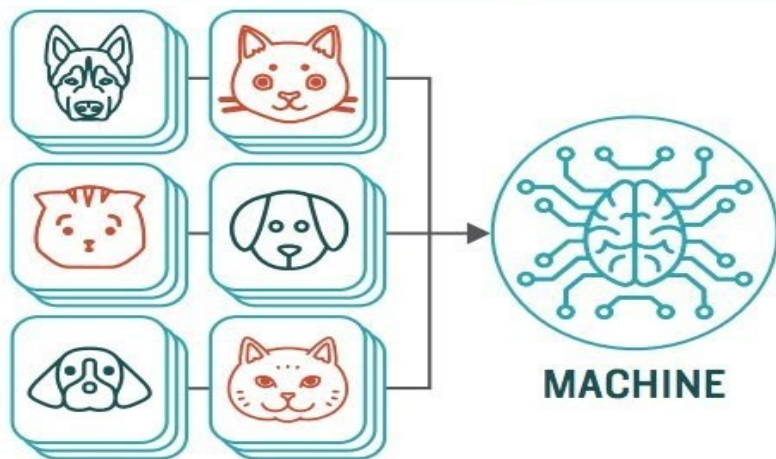
- Unsupervised learning (UL) is a type of algorithm that learns patterns from untagged data.
- The hope is that, through mimicry, the machine is forced to build a compact internal representation of its world and then generate imaginative content.
- In contrast to supervised learning (SL) where data is tagged by a human, e.g. as "car" or "fish" etc, UL exhibits self-organization that captures patterns as neuronal predilections or probability densities.

Unsupervised Machine Learning

How **Unsupervised** Machine Learning Works

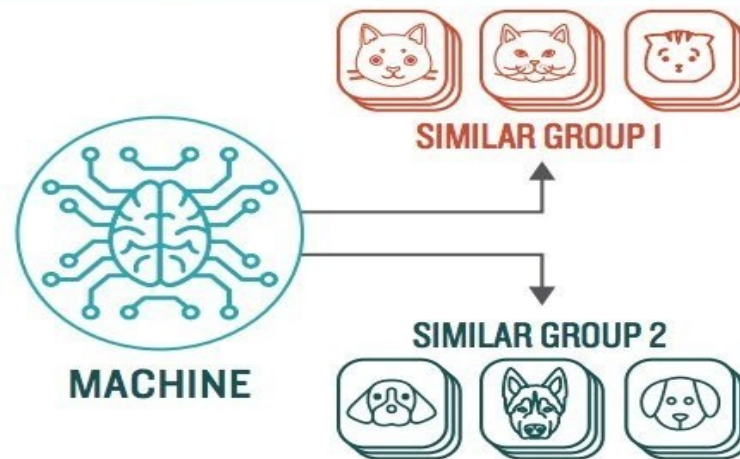
STEP 1

Provide the machine learning algorithm uncategorized, unlabeled input data to see what patterns it finds



STEP 2

Observe and learn from the patterns the machine identifies

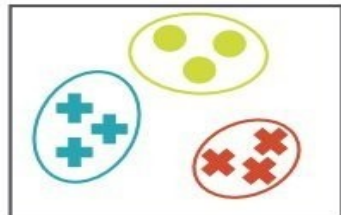


TYPES OF PROBLEMS TO WHICH IT'S SUITED

CLUSTERING

Identifying similarities in groups

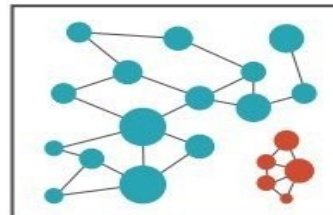
For Example: Are there patterns in the data to indicate certain patients will respond better to this treatment than others?



ANOMALY DETECTION

Identifying abnormalities in data

For Example: Is a hacker intruding in our network?



Un-Supervised Learning : Examples

- Clustering methods include: hierarchical clustering, k-means, mixture models, DBSCAN, and OPTICS algorithm
- Anomaly detection methods include: Local Outlier Factor, and Isolation Forest
- Learning latent variable models such as Expectation–maximization algorithm (EM), Method of moments, and Blind signal separation techniques (Principal component analysis, Independent component analysis, Non-negative matrix factorization, Singular value decomposition)

Reinforcement Learning

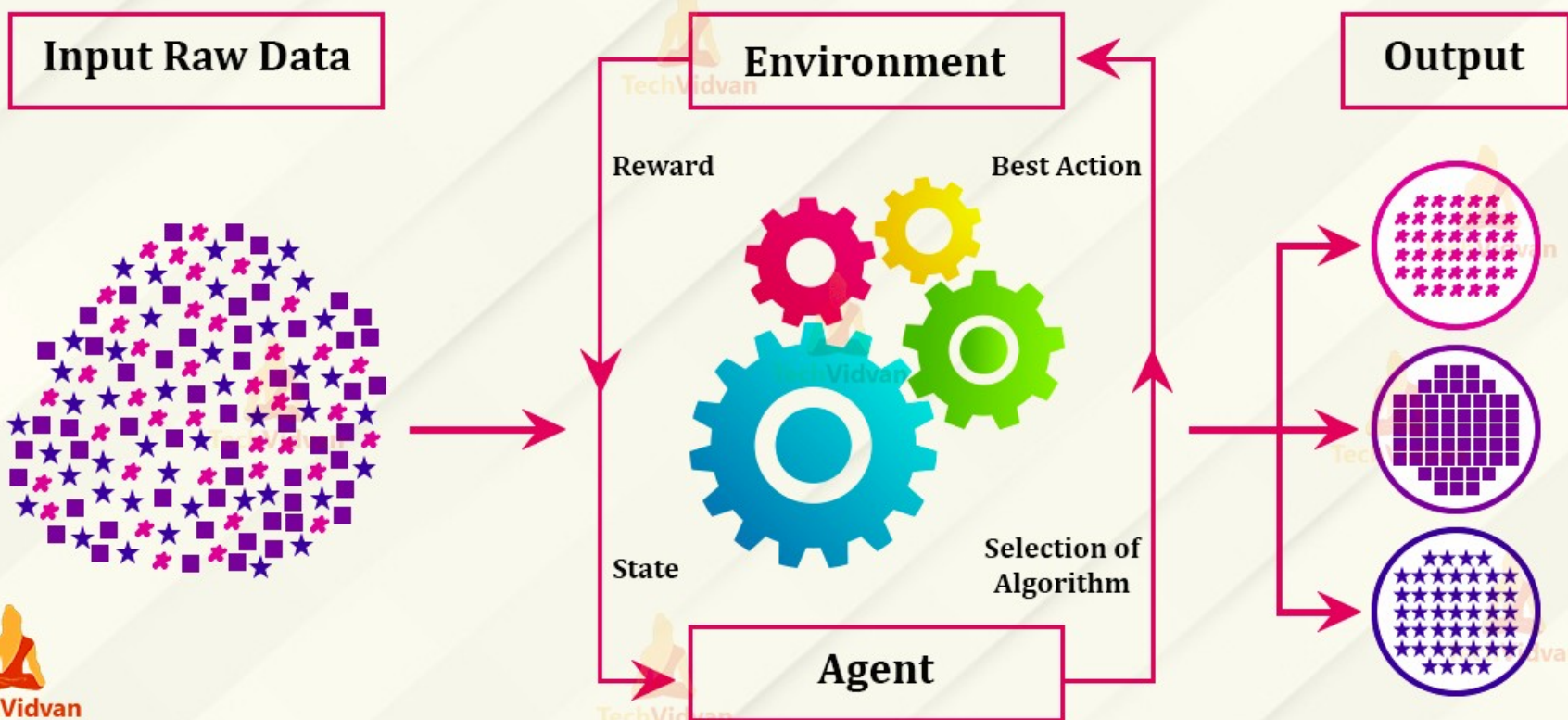
- Reinforcement Learning is defined as a Machine Learning method that is concerned with how software agents should take actions in an environment.
- Reinforcement Learning is a part of the deep learning method that helps you to maximize some portion of the cumulative reward.

Reinforcement Learning

- Imagine someone playing a video game. The player is the agent, and the game is the environment. The rewards the player gets (i.e. beat an enemy, complete a level), or doesn't get (i.e. step into a trap, lose a fight) will teach him how to be a better player.
- In supervised learning, for example, each decision taken by the model is independent, and doesn't affect what we see in the future.
- In reinforcement learning, instead, we are interested in a long term strategy for our agent, which might include sub-optimal decisions at intermediate steps, and a trade-off between exploration (of unknown paths), and exploitation of what we already know about the environment.

Reinforcement Learning

Reinforcement Learning in ML



Conclusion

A breakthrough in
machine learning would
be worth ten Microsofts.

Bill Gates

quote fancy

Useful web resources

- www.mitu.co.in
- www.pythonprogramminglanguage.com
- www.scikit-learn.org
- www.towardsdatascience.com
- www.medium.com
- www.analyticsvidhya.com
- www.kaggle.com
- www.stephacking.com
- www.github.com

Thank you

This presentation is created using LibreOffice Impress 7.4.1.2, can be used freely as per GNU General Public License



@mitu_skillologies



@mITuSkillologies



@mitu_group



@mitu-skillologies



@MITUSkillologies

kaggle

@mituskillologies

Web Resources
<https://mitu.co.in>
<http://tusharkute.com>



@mituskillologies

contact@mitu.co.in
tushar@tusharkute.com