Introduction to Artificial Intelligence

By Neil Bloomberg

Bowie Computer Club
Spring 2019

Introduction

This presentation will cover the fundamentals of Artificial Intelligence, how machines learn, and how this learning is being used today.

Disclaimer → I am not an expert in this field

Agenda

- Definition of artificial intelligence
- How the human brain works
- Concepts of learning
- Neural networks
- Computers versus humans
- Applications and where are we heading
- Q & A

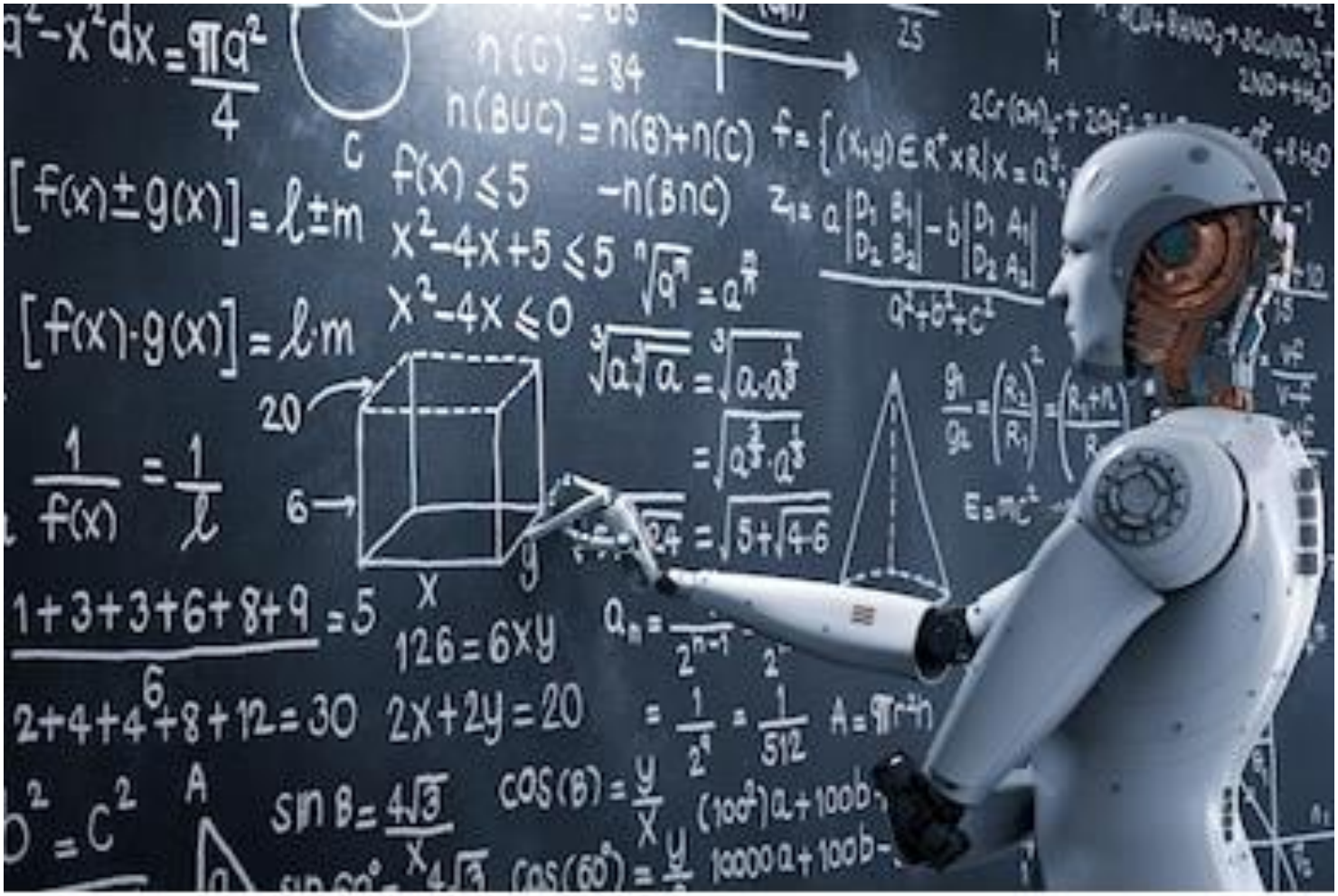
Acknowledgements etc.

→All clip art used in this presentation was obtained from Wikipedia unless otherwise noted

→ Special thanks to the book “Thinking Machines - The Quest for AI” by Luke Dormehl

→ Slide take-away points are designed with an Arrow and highlighted at the bottom in **bold RED**

Artificial Intelligence - Welcome!



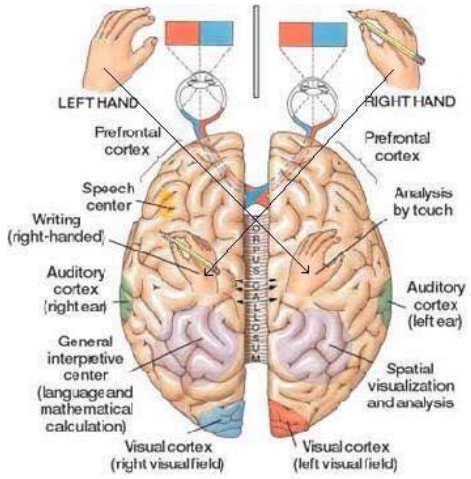
Definition of Artificial Intelligence

Artificial Intelligence (AI) is the theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages. **Source: Wikipedia**

Artificial Intelligence (AI) - Computer systems able to learn from data received or able to perform tasks normally associated with human thinking. **Source: Neil**

How the human brain works

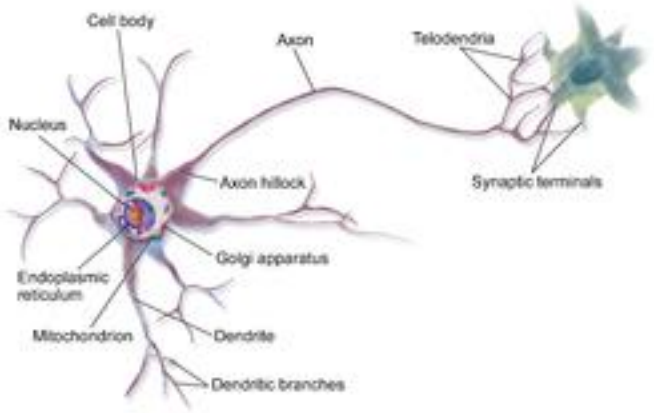
The **brain** contains about 100 billion nerve cells arranged in patterns that coordinate thought, emotion, behavior, movement, and sensation. While all the parts of your brain work together, each part is responsible for a specific function - controlling everything from your heart rate to your mood.



Brain functionality

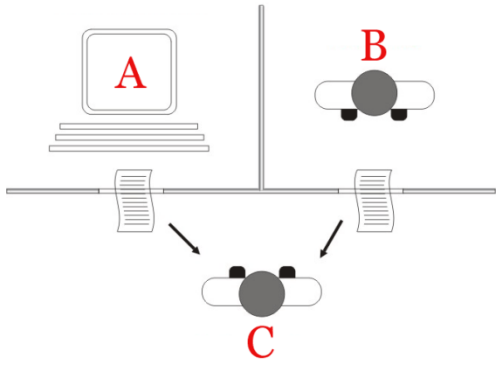
The **neuron** is the basic working unit of the brain, a specialized type of cell designed to transmit information to other neurons as well as nerve cells, muscle cells, or gland cells. Neurons and their connective synapses are used by the brain for:

- Sensory interpretation - Sight, sound, taste, touch, and hearing
- Control outputs - Muscle movement, breathing, heart rate, and other motor functions
- Memory - Storing information
- Thinking - How to do tasks, thinking, and learning



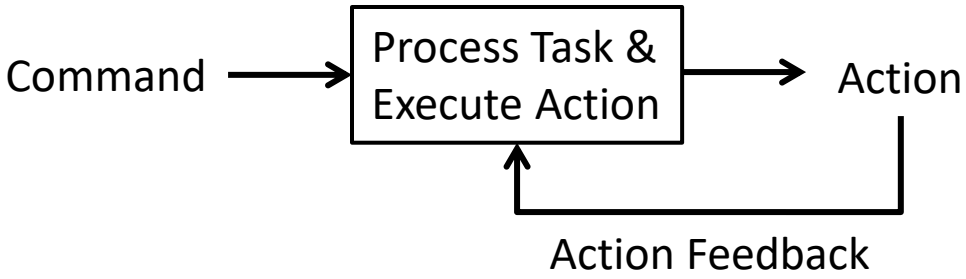
Turing test

The **Turing test**, developed by Alan Turing in 1950, is a test of a machine's ability to exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human. Turing proposed that a human evaluator would judge natural language conversations between a human and a machine designed to generate human-like responses. The evaluator would be aware that one of the two partners in conversation is a machine and all participants would be separated from one another.



Computer learning process

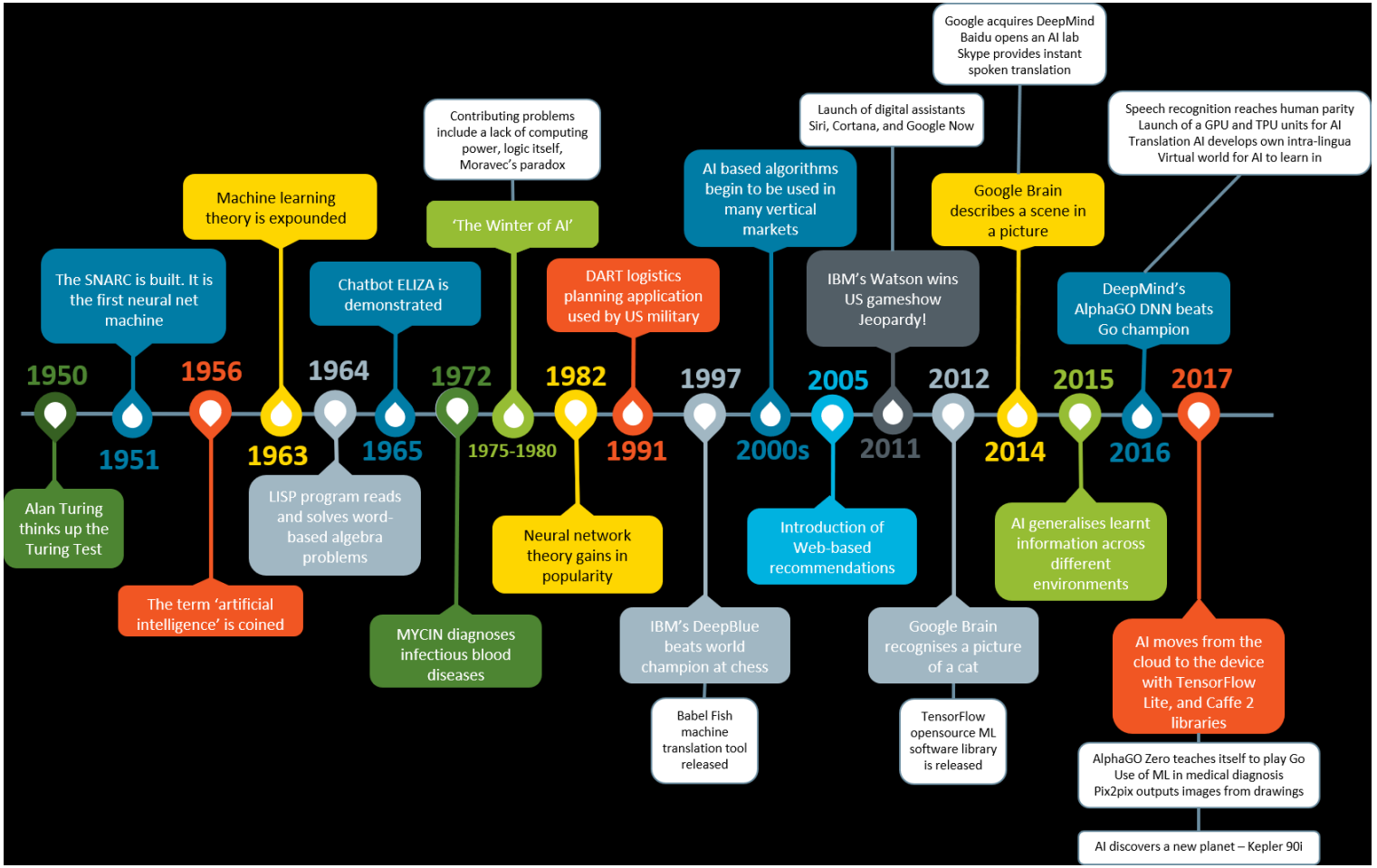
Most computer models use statistical algorithms that receive a single command, analyze the command (understand), take some action(perform work), and measure the results as success or failure. The results are then used to adjust the statistical formula values until the desired result is consistently achieved.



→ This is the basic model for all computer learning is derived

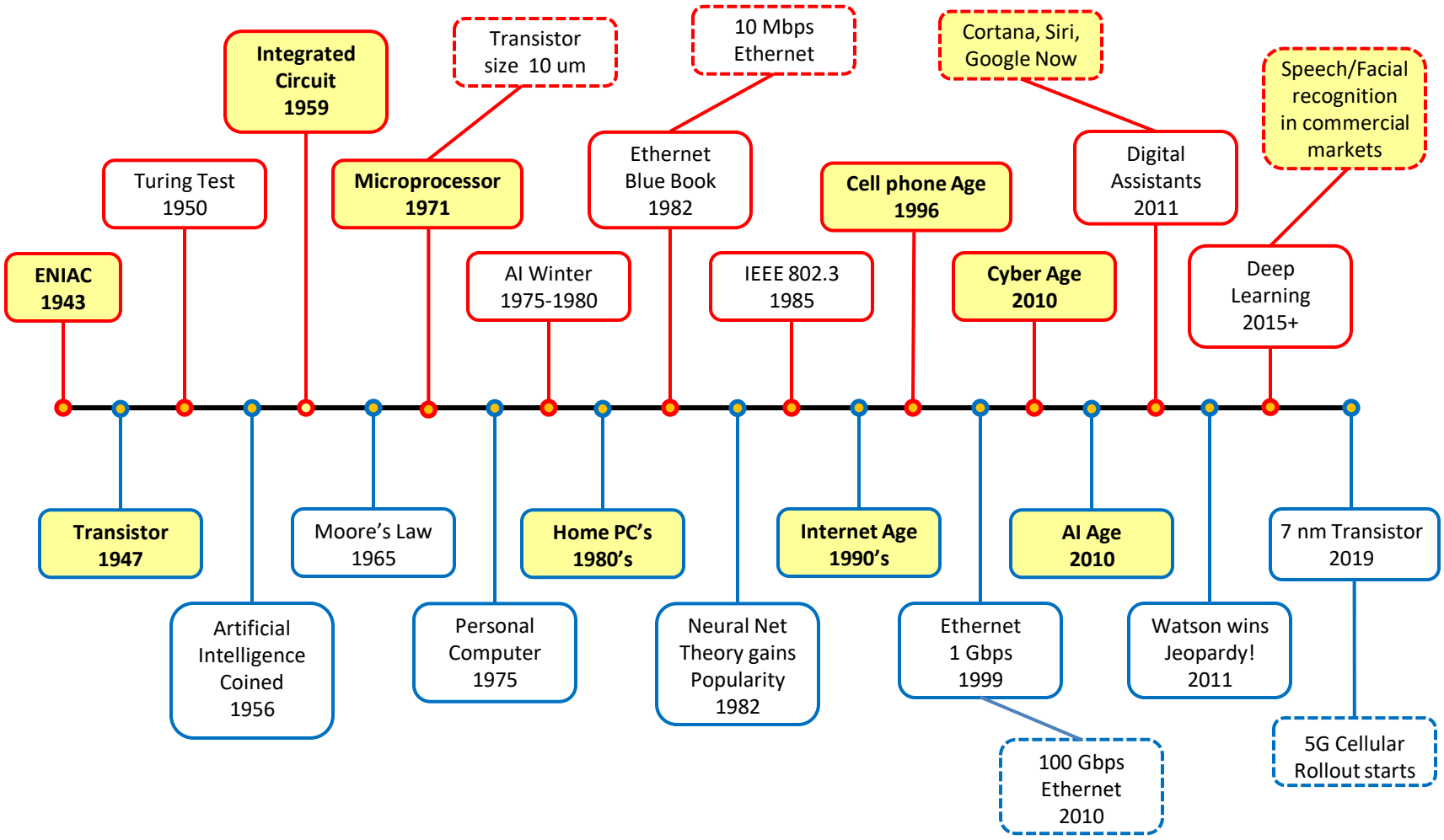
History of AI - Timeline

1947 – ENIAC
1947 – Transistor
1959 – I.C.
1950 – Turing Test
1956 – AI coined
1965 – Moore's Law
1971 – Intel 4004 uP
1975 – P.C.'s
1980's – Home Computers
1982 – Ethernet
Blue Book
1982 – Neural Net Theory
1985 – IEEE 802.3
1990's – Internet
1999 – 1 Gbps
2000's – 100 Gbps
2010 – Cyber Age
2010 – AI Age
2011 – Digital Assistants
2011 – Watson wins Jeopardy!
2015+ Speech/facial Recognition
2020 – 1 Tbps…



→ 2010 AI began a new type of Technical Revolution comparable to the Industrial Revolution

History of computer technology



AI statistical algorithms

Machine learning algorithms build statistical models of sample data, known as training data, in order to make predictions or decisions without being explicitly programmed to perform the task.

- **Supervised** learning algorithms take numerous data sets containing inputs along with the desired outputs to train the algorithms before inputs without known outputs are introduced.

- **Unsupervised** learning algorithms take a set of data that contains only inputs and find structure in the data (e.g. by grouping or clustering data points). The algorithms therefore learn from test data that has not been labeled, classified, or categorized.

- **Reinforcement** learning is concerned with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward (e.g. video game score).

Deep Learning

Deep learning, especially deep neural networks, is a form of machine learning largely responsible for recent explosions of AI success rates. Deep learning may be used for all types of machine learning:

- Top image classification (supervised learning) successes use a specialized type of deep learning called convolutional neural networks (CNNs)
- Deep anomaly detection is an example of unsupervised learning that can use another specialized form of deep learning called recurrent neural networks (RNNs)
- Deep reinforcement leaning can use a variety of deep learning algorithms to make decisions that optimize rewards in a given environment (e.g. CNN processing of video game images)

Layers of deep neural networks can learn feature representation that sometimes have intuitive understandable interpretations. For example, image recognition.

Deep learning - Image recognition

Deep learning image recognition (unsupervised) has networks capable of learning from data that is unstructured or unlabeled using a layered approach. Example from image recognition:

- A facial image is collected consisting of a massive set of pixels
- The pixels are used to calculate lines and boarders during the next layer
- The next layer processes contours within the lines and boarders
- The next layer identifies different shades within the regions
- The next layer of processing identifies any unusual regions

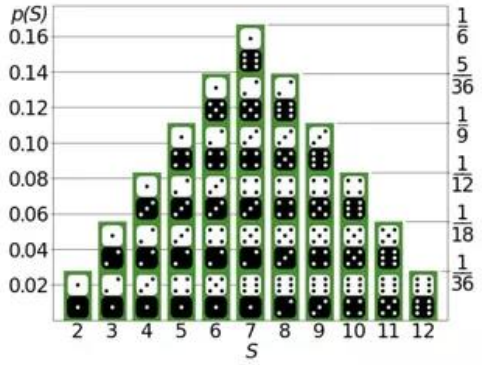
Using large data sets of millions of facial images, the program is able to distinguish the many minor differences between faces and attach labels (**tagging**) to them such that the same face at any angle or lighting condition will result in a perfect image match.

AI - A look inside

The math behind AI

Most AI products incorporate the following into their models:

- Probability and statistics
- Linear algebra
- Optimization theory
- Calculus
- Matrix theory



$$V_{ix}\left(t+\frac{\Delta t}{2}\right)=V_{ix}\left(t-\frac{\Delta t}{2}\right)+\frac{\Delta t}{M_i}R_{ix}(t)$$

$$X_i(t+\Delta t)=X_i(t)+\Delta t\times V_{ix}\left(t+\frac{\Delta t}{2}\right)$$

Where:

Δt is the time interval between two updates.

By the principle of equilibrium of forces, the relationship between the residuals and the geometry can be obtained:

$$R_{ix}(t+\Delta t)=P_{ix}(t+\Delta t)+\sum\frac{T_m(t+\Delta t)}{l_m(t+\Delta t)}\times(X_j(t+\Delta t)-X_i(t+\Delta t))$$

AI statistical models

In short, it is all about math. Machine learning algorithms are described as learning a target function (f) that best maps input variables (X) to an output variable (Y): **Y = f(X).**

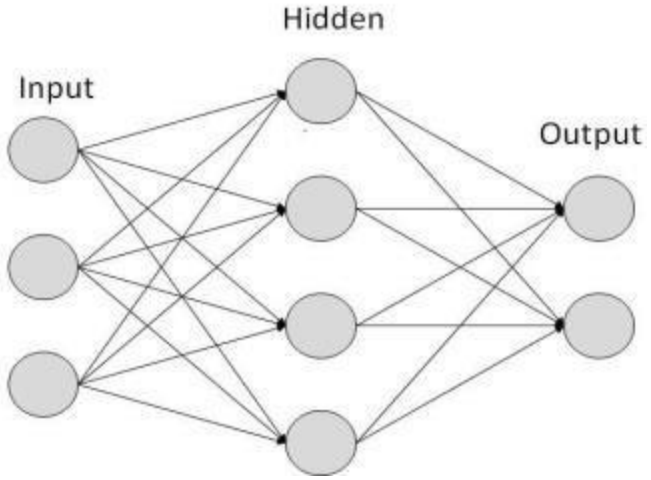
Just like there are numerous types of neurons, there are also numerous mathematical models for the processing of information. Below is just a sample of some algorithm variety used in the field of AI:

- Linear Regression
- Logical Regression
- Linear Discriminant Analysis
- Classification and Regression Trees
- Naive Bayes
- K-Nearest Neighbors
- Learning Vector Quantization
- Support Vector Machines
- Bagging and Random Forest
- Boosting and AdaBoost

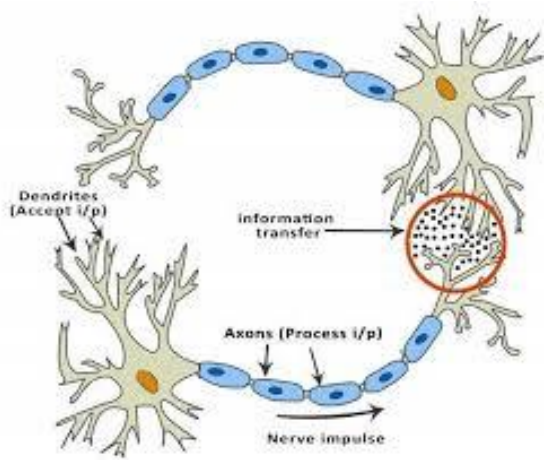
→ This is just a small sample of the growing AI algorithms in use today

Neural networks

An artificial **neural network** is an interconnected group of computer nodes, similar to the vast network of neurons in a brain. Here, each circular node represents an artificial neuron and an arrow represents a connection from the output of one artificial neuron to the input of another.



Computer Neural Network



Human Brain Neural Network

→ By reverse engineering how the human brain works, scientists and engineers are working toward creating an AI device that performs the same functions

Search engines vs do engines

Search engines, like Google, are applications that when given a query will perform a world wide web search based upon the words entered into the query. The result is a list of the most likely word matches, which can be quite extensive.

Do engines, also known as AI Assistants, apply methods of understanding to the question then perform work on the vast amounts of data available resulting in a single answer. (e.g. Siri, Cortana, Google Now).

→ Search engines are not typically classified as AI software

Computers versus humans

The field of Artificial Intelligence has made significant progress in the past 40 years, with computers able to:

- **Facial recognition** - The ability to distinguish a single individual from millions of photos
- **Voice recognition** - The ability to distinguish a single individual from millions of voices and sounds
- **Language translation** - (e.g. Google Translate) The ability to translate one spoken language to another using the speakers own tonal voice as an output
- **Quickly learn** – Strategy optimization with the ability to win at most video games

AI applications

The field of Artificial Intelligence has opened up many opportunities:

- **Smart drones** - Ability of a drone (UAV) to independently take off and land between two distant areas. In flight autonomous refueling, collision avoidance, and take counter measures against assault

- **Medical applications** - With the ability to draw upon large quantities of medical data, AI medical applications can help doctors to diagnose patients and provide assistance to create new drugs

- **Automotive** - Smart Cars are now alerting drivers as to road hazards with self driving cars/trucks already in use in small quantities
  (Note →still 10 years away)

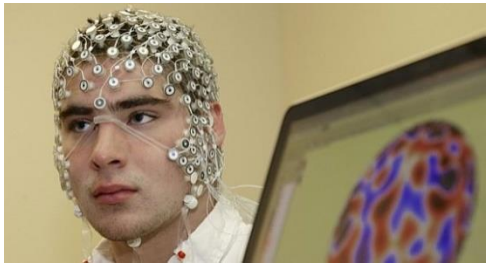
- **Robotics** - Smart robots are now performing tasks such as cleaning, machine assembly, and delivering packages within buildings

AI and Drones

"Military and Aerospace" Magazine, 18 Sept. 2018 - A person with a brain chip can now pilot a swarm of drones - or even advanced fighter jets, thanks to research funded by the U.S. military's Defense Advanced Research Projects Agency (DARPA).

The work builds on research from 2015, which allowed a paralyzed woman to steer a virtual F-35 Joint Strike Fighter with only a small, surgically-implantable microchip. Agency officials announced that they had scaled up the drone swarm technology to allow a user to steer multiple jets at once.

More importantly, DARPA was able to improve the interaction between pilot and the simulated jet to allow the operator, a paralyzed man, to not only send but receive signals from the craft.



AI Assistants – Legal profession

LegalZoom is one of the leading legal AI Assistants used for wills, trusts, estate planning, legal research, patent research, and due diligence in the legal industry.

Kira is a powerful AI learning software that identifies, extracts, and analyzes text in contracts and other documents.

Luminance is the market-leading AI platform for legal due diligence in the legal industry. Luminance understands language the way humans do, in volumes and at speeds that humans will never achieve. It provides an immediate and global overview of any company, picking out warning signs without needing any instruction. Whether used for due diligence, compliance, insurance, or contract management, Luminance adds value to a legal team, freeing lawyers to focus on what matters.

rradar AI is an AI Assistant for legal advice, tools and training to spot, manage, and minimize business risks; educating businesses to stay on the right side of the law and guiding it to safety if legal problems do arise.

→ Like TurboTax, Legal AI software is becoming more commonplace and prices are falling

AI Assistants – Medical profession

Medical AI virtual Assistants operating in response to voice or text interactions have steadily gained traction and formed a profitable sector, according to *Research and Markets.* Market research firms project that total revenue will hit $15.8 billion in 2021 up from $1.6 billion in 2015. It is also estimated that total global consumers will reach 1.8 billion by 2021.

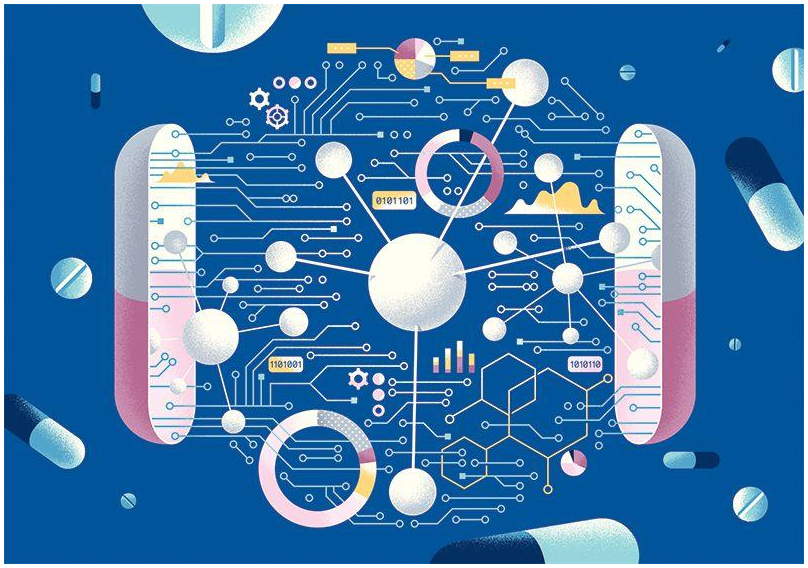
Research suggests that the majority of AI use cases and emerging applications for virtual medical assistants appear to fall into four main categories:

- **Medical record navigation -** When companies and medical professionals use machine learning and natural language processing to search, analyze, and record clinical data in a patient's electronic health record (EHR)
- **Medical transcription** - When medical professionals use machine learning and natural language processing to transcribe clinical data recorded during patient visits
- **Medical information search** - When companies develop machine learning algorithms for chatbots to deliver personalized responses to searches for medical information by consumers and healthcare professionals based on consumer or patient data
- **New drug discovery -** Machine learning and other technologies are expected to make the hunt for new pharmaceuticals quicker, cheaper, and more effective

Drug discovery

An enormous figure looms over scientists searching for new drugs: the estimated US $2.6-billion price tag of developing a treatment. A lot of that effectively goes down the drain, because it includes money spent on the nine out of ten candidate therapies that fail somewhere between phase I trials and regulatory approval. Few people in the field doubt the need to do things differently. Leading biopharmaceutical companies believe a solution is at hand. Pfizer is using **IBM Watson**, a system that uses machine learning, to power its search for immuno-oncology drugs.

Sanofi has signed a deal to use UK start-up **Exscientia's** Artificial Intelligence platform to hunt for metabolic-disease therapies, and Roche subsidiary Genentech is using an AI system from GNS Healthcare in Cambridge, Massachusetts, to help drive the multinational company's search for cancer treatments. Most sizeable biopharma players have similar collaborations or internal programs.



→ AI software may be able to reduce these costs significantly

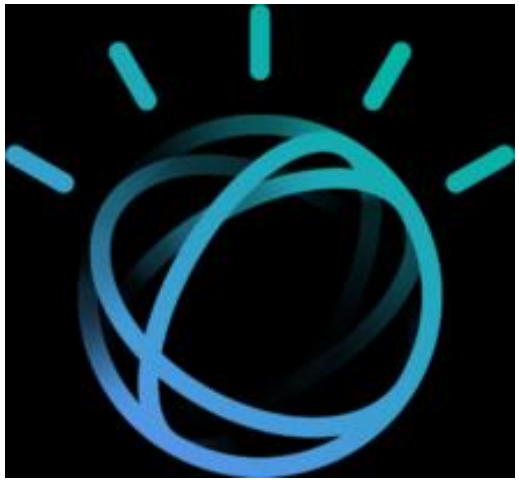
Autonomous cars

The 2017 fatality caused by an Uber self driving car has caused auto manufactures to rethink safety and address numerous concerns:

- Reduced visibility - detecting lane markers and seeing past snow, rain, fog, or sand storms
- Snow and rain driving - slippery road, ice patches, recovery if car goes into spin
- Driving on roads without visible lane markers
- Making left hand turns
- Traffic obstructions
- Driving along with humans who make sudden mistakes
- Sensor failures are common
- ISO 26262 Safety Standards
- Legal issues still being figured out

IBM Watson project

Created in the 1990's, Watson was first designed to play the game of jeopardy pitted against two leading Jeopardy winners: Ken Jennings (with a record winning streak of 73 games and over 2 million dollars in winnings) and Brad Rutter. Watson won!



Microsoft Tay project

Tay was an artificial Intelligence chatter bot that was originally released by Microsoft Corporation via Twitter on March 23, 2016; it caused subsequent controversy when the bot began to post inflammatory and offensive tweets through its Twitter account, forcing Microsoft to shut down the service only 16 hours after its launch. According to Microsoft, this was caused by trolls (bad guys) who "attacked" the service as the bot made replies based on its interactions (learning) with people on Twitter. It was soon replaced with **Zo**, another AI chatter bot with filters based upon lessons learned from the Tay project.

→ Sometimes AI projects can go wrong

Challenges for AI

Scientists and researchers are trying to pattern AI after the human brain. With over 100 billion neurons in a human brain and over 1,000 specialized neuron types, scientists still do not understand exactly how neurons process information or how they communicate with each other. The current focus is on reverse engineering small animals (e.g. mice) to gain insight into how the brain works.

→ With billions of dollars being spent each year in a multiple fields, AI research is already paying off

AI limitations

AI application may be good at facial recognition, language translation, and many other tasks, but much more is needed in the following areas:

- Creativity
  - Creating new forms of art that express beauty and emotion
  - Creating something new that seems illogical

- Speech Recognition and understanding
  - Understanding puns, oblique references, satire, and jokes

- General
  - Making moral judgements of right and wrong

Possible job industries at risk

As progress is being made in the field of AI, many low level tasks will soon be replaced by automated processes. These include, but are not limited to following job industries:

- Insurance Underwriters and Claims Representatives. The effects of automation on the insurance industry are already being felt
- Bank Tellers and Representatives
- Financial analysts
- Construction workers
- Farmers
- Taxi drivers
- Manufacturing workers

AI in movies - fact or fiction

In the 1982 action movie, Firefox, actor Clint Eastwood plays the role of an American fighter pilot who is fluent in Russian. As an ace pilot he is assigned to steal an advanced Soviet fighter jet that has the ability to be controlled telepathically as well as makes decisions on its own. Sound familiar? Like Star Trek, what was once portrayed as science fiction is now becoming a reality:

- Thought control – small electrodes placed on a pilots head can now control a fleet of drones. In the commercial sector, a few electrodes on a head band can be used to turn on home appliances including lights and change television channels
- Cell phones – The Star Trek communicator appearing in the original TV series in 1965 has not only become a reality but is now considered old technology

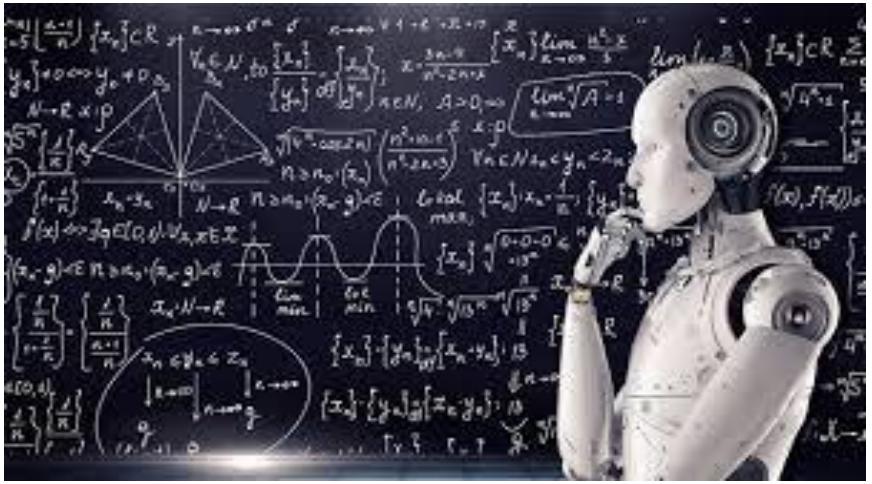
It is amazing to see the number of futuristic devices portrayed in movies long ago becoming a reality. However, if you're thinking about time travel, well that will not happen!

Summary

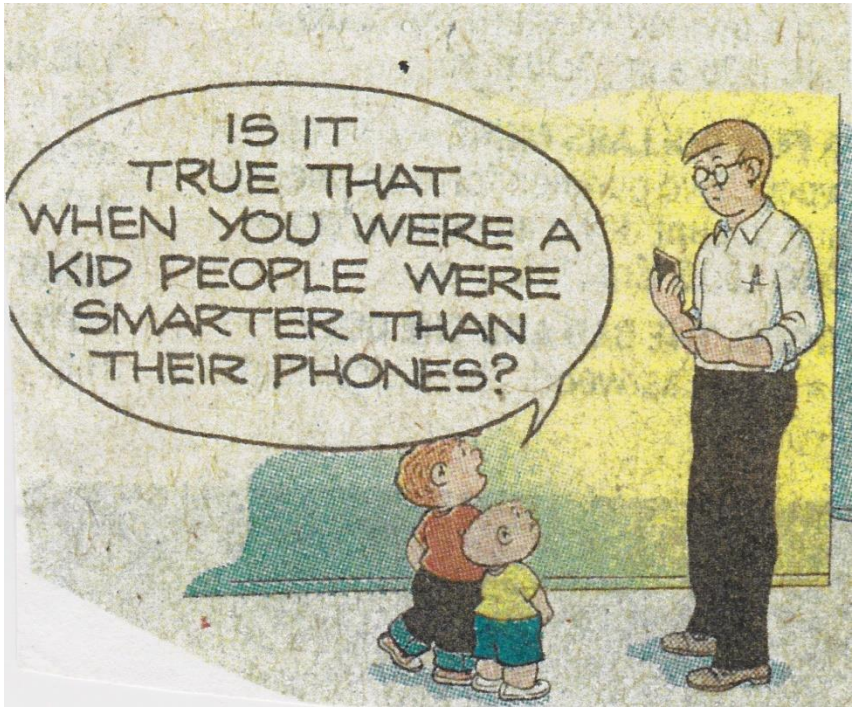
The past 60 years have shown great progress in the field of AI. Speech and facial recognition are now common on most smart phones. Language translation is becoming easier (e.g. Google Translate). Self driving cars are still ten years away from large scale use.

Around 1760, the Industrial Revolution transformed the West from an agricultural society to an industrial society, replacing many human labor activities with machines. We are now entering a new era in which Artificial Intelligence will continue this same trend in a more profound way with the evolution of AI Assistants:

- Secretaries
- Phone answering services
- Legal researchers (LegalZoom)



Q & A ?????


IS IT TRUE THAT WHEN YOU WERE A KID PEOPLE WERE SMARTER THAN THEIR PHONES?

Backup Slides

Supervised learning

Supervised learning, as the name indicates, involves the presence of a supervisor or teacher. Basically supervised learning is a learning in which we train and then verify the machine results starting with input data that is already tagged with the correct output answer. Each time new data is provided the machine output is assessed. After a sufficient number of samples with known outputs are given, new data without a known output is provided until the machine can correctly provide a correct output consistently.

For instance, suppose you are given an basket filled with different kinds of fruits. Now the first step is to train the machine with all different fruits one by one like this:

If the object is red, round, and depression at top then it will be labelled as an **Apple**.
If the object is green-yellow with a long curving cylinder shape then it will be labelled as a **Banana**.

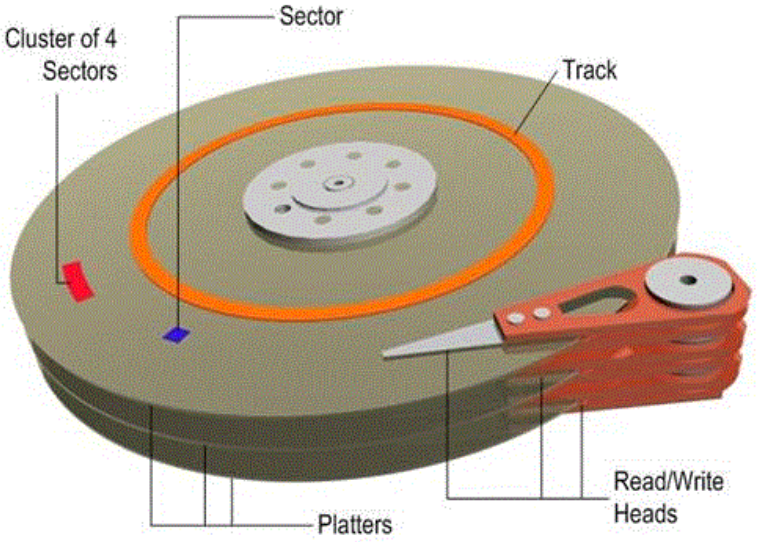
Shamir Algorithm

In 2013, Assistant Professor Lior Shamir from the Lawrence Institute of Technology, Southfield, MI, developed audio analysis technology to study the vocal communication of whales, and they expanded the algorithm to analyze the albums of the Beatles and other well-known bands such as Queen, U2, ABBA and Tears for Fears.

The algorithm works by first converting each song to a spectrogram – a visual representation of the audio content. That turns an audio analysis task into an image analysis problem, which is solved by applying comprehensive algorithms that turn each music spectrogram into a set of almost 3,000 numeric descriptors reflecting visual aspects such as textures, shapes and the statistical distribution of the pixels. Pattern recognition and statistical methods are then used to detect and quantify the similarities between different pieces of music.

Case Study – AI Caching Disk Controller

In 1986, U.S. Design Corp, a subsidiary of Maxtor Corporation, created a AI based technology to reduce access time to magnetic disk drives. Currently, disk access time average was about 23 milliseconds (move disk head to proper track and locate proper sector on the track. Using historical information of disk head movement, read ahead buffering, and sequencing the writing of data, disk head movement was substantially minimized to an average access time of 50 microseconds. Benchmark studies using a simple Apple microcomputer and the caching disk technology resulted in the Apple computer out performing many Mini computers.



Hard Disk Drive 3D Image - Sourced From TechNet, Microsoft Website

Computer learning

“Artificial Intelligence is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment.” Source Wikipedia

AI Assistants - Medical Profession

MedWhat, is a medical virtual assistant that answers medical and health questions for consumers and doctors instantly. The answers are provided by an intelligent super-computer that learns about medicine everyday and over time about your health record and medical questions history.

Faima also a medical virtual assistant uses all the benefits of the Artificial Intelligence to diagnose diseases and provide users with the best treatment possible. This project is not setting any limits: Faima is available all over the world – and that may become the key to its success. It is needless to say that the earlier the patient gets diagnosed, the easier it is to cure the illness. The Artificial Intelligence is used to analyze each exact case based on the data from more than forty million patients.

AI Assistants - Legal Profession

Vincent, developed by vLex, the first AI-powered intelligent legal research assistant of its kind. Only Vincent can analyze documents in two languages (English and Spanish) from 9 countries (and counting), and is built ready to incorporate content not only from vLex's expansive global collection, but also from internal knowledge management resources, public sources and licensed databases *simultaneously*. How does Vincent do it, you ask? Well, it's been trained on vLex's extensive global collection of 100 million+ legal documents

Ailira is an Artificial Intelligence developed by Cartland Law that uses natural language processing to provide free legal information on a broad range of legal issues, including Business Structuring, Wills and Estate Planning and much more coming soon! In addition, you can use Ailira to instantly generate Australian legal documents for your business and personal use, much cheaper and faster than a visit to a lawyer would take.

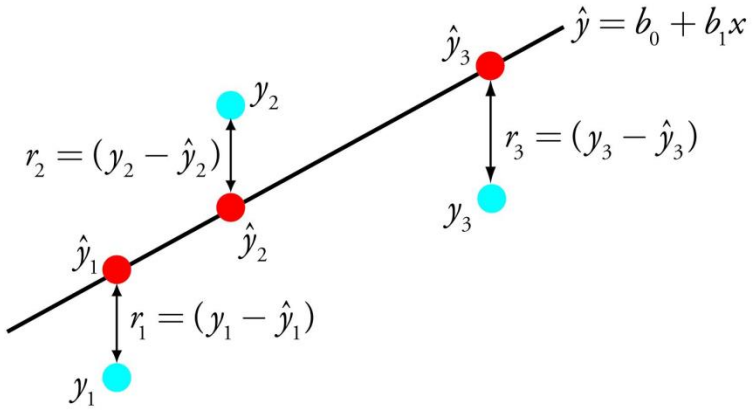
→ AI assistants in the legal profession are performing many of the legal research once performed by entry level lawyers and paraprofessionals.

Example - Linear Regression

Linear regression is perhaps one of the most well-known and well-understood algorithms in statistics and machine learning.

Predictive modeling is primarily concerned with minimizing the error of a model or making the most accurate predictions possible, at the expense of explainability. We will borrow, reuse and steal algorithms from many different fields, including statistics and use them towards these ends.

The representation of linear regression is an equation that describes a line that best fits the relationship between the input variables (x) and the output variables (y), by finding specific weightings for the input variables called coefficients (B).

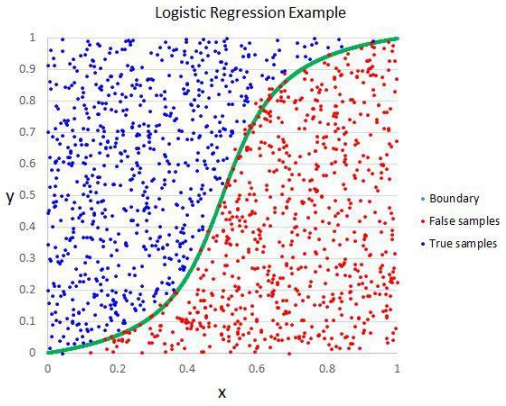


Logistic Regression

Logistic regression is another technique borrowed by machine learning from the field of statistics. It is the go-to method for binary classification problems (problems with two class values).

Logistic regression is like linear regression in that the goal is to find the values for the coefficients that weight each input variable. Unlike linear regression, the prediction for the output is transformed using a non-linear function called the logistic function.

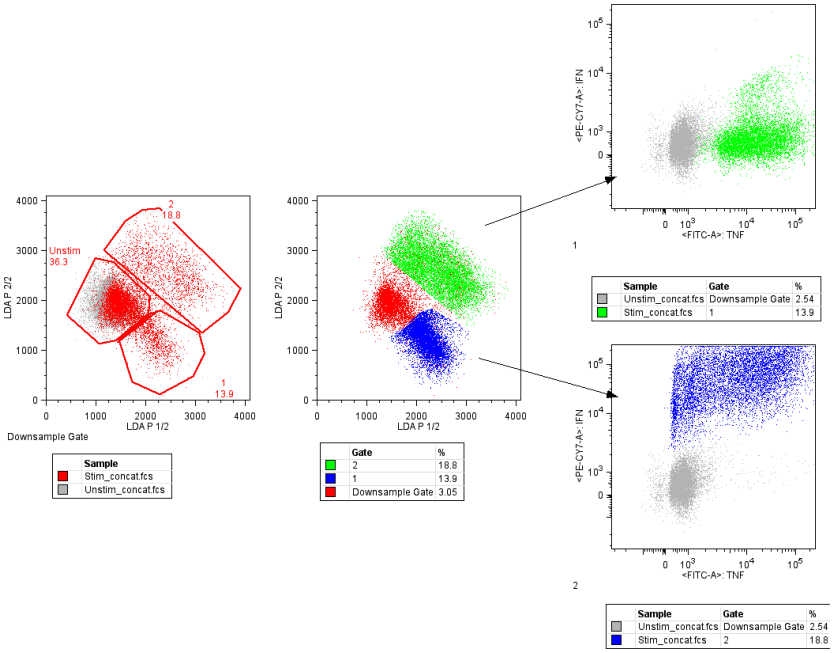
The logistic function looks like a big S and will transform any value into the range 0 to 1. This is useful because we can apply a rule to the output of the logistic function to snap values to 0 and 1 (e.g. IF less than 0.5 then output 1) and predict a class value.



Linear Discriminant Analysis

Logistic Regression is a classification algorithm traditionally limited to only two-class classification problems. If you have more than two classes then the Linear Discriminant Analysis algorithm is the preferred linear classification technique. The representation of LDA is pretty straight forward. It consists of statistical properties of your data, calculated for each class. For a single input variable this includes:

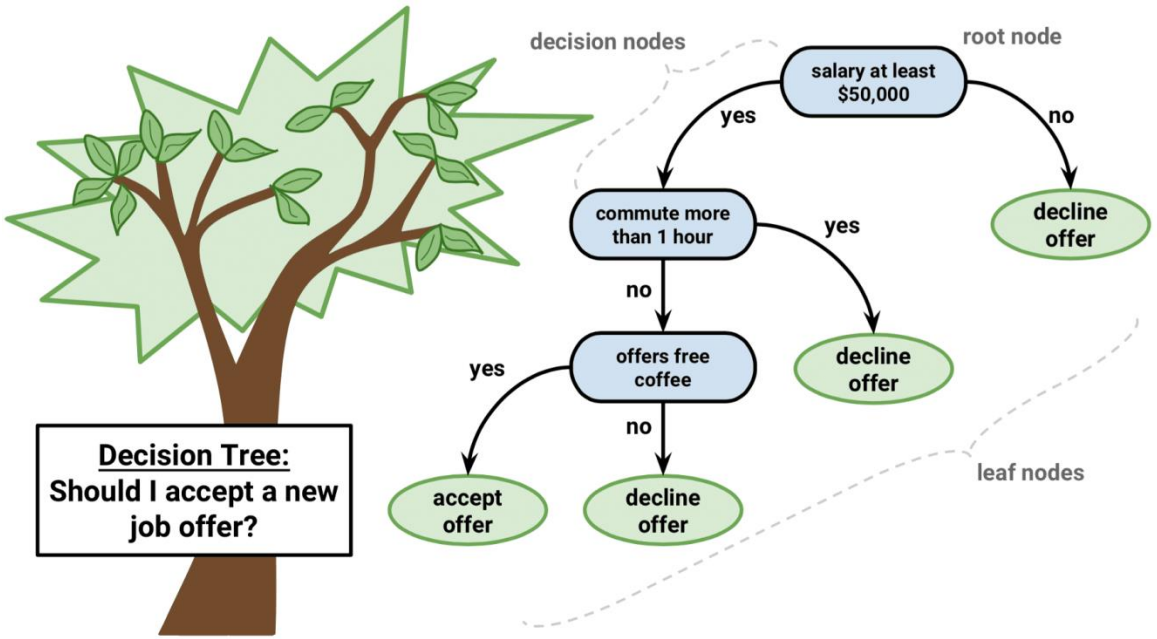
- The mean value for each class
- The variance calculated across all classes



Classification and Regression Trees

Decision Trees are an important type of algorithm for predictive modeling machine learning.

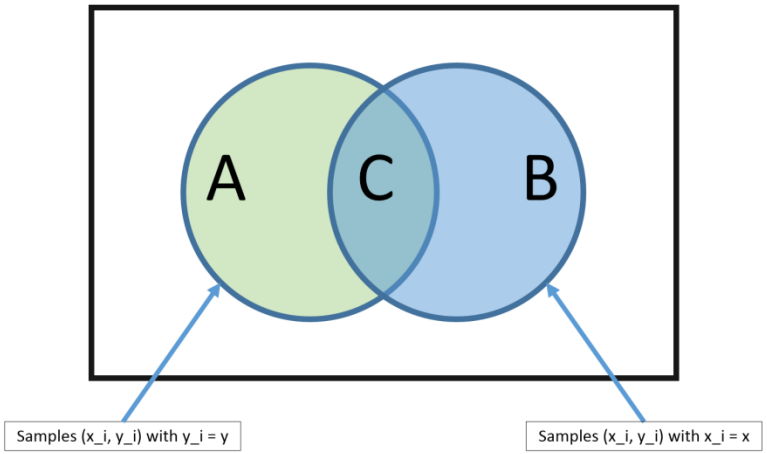
The representation of the decision tree model is a binary tree. This is your binary tree from algorithms and data structures, nothing too fancy. Each node represents a single input variable (x) and a split point on that variable (assuming the variable is numeric).



Naive Bayes

Naive Bayes is a simple but surprisingly powerful algorithm for predictive modeling.

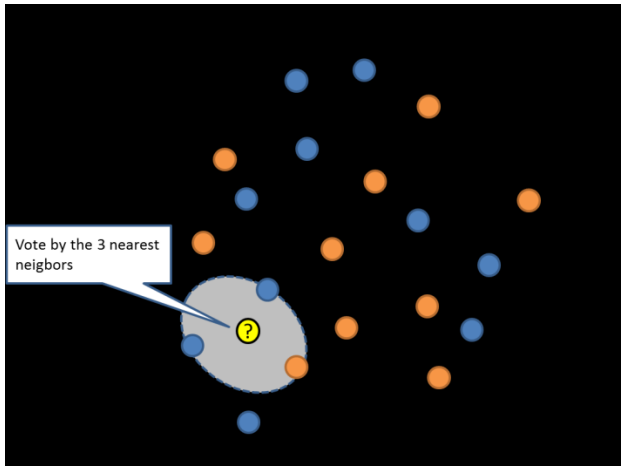
The model is comprised of two types of probabilities that can be calculated directly from your training data: 1) The probability of each class; and 2) The conditional probability for each class given each x value. Once calculated, the probability model can be used to make predictions for new data using Bayes Theorem. When your data is real-valued it is common to assume a Gaussian distribution (bell curve) so that you can easily estimate these probabilities.



K-Nearest Neighbors

The KNN algorithm is very simple and very effective. The model representation for KNN is the entire training dataset. Simple right?

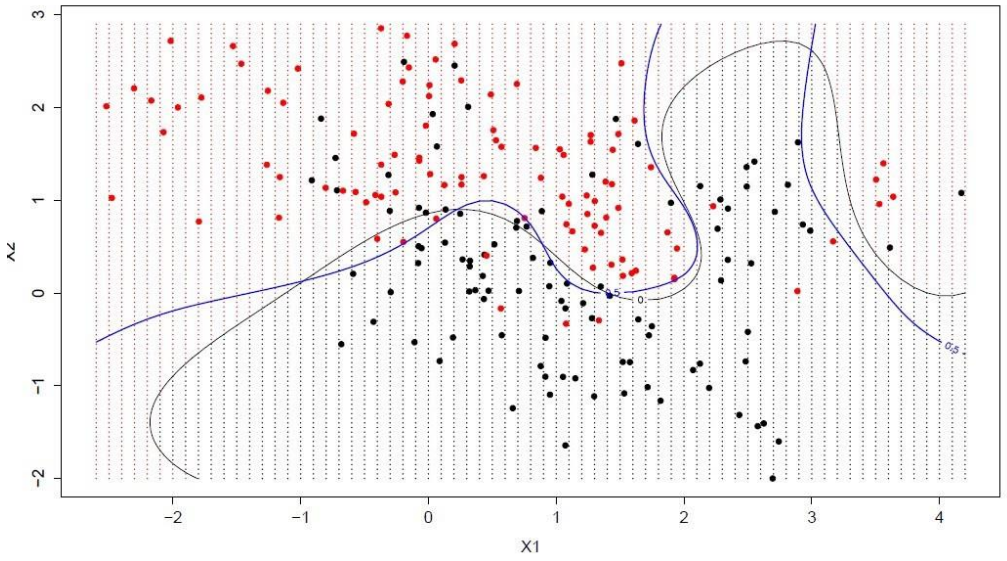
Predictions are made for a new data point by searching through the entire training set for the K most similar instances (the neighbors) and summarizing the output variable for those K instances. For regression problems, this might be the mean output variable, for classification problems this might be the mode (or most common) class value. The trick is in how to determine the similarity between the data instances. The simplest technique if your attributes are all of the same scale (all in inches for example) is to use the Euclidean distance, a number you can calculate directly based on the differences between each input variable.



Support Vector Machines

Support Vector Machines are perhaps one of the most popular and talked about machine learning algorithms.

A hyperplane is a line that splits the input variable space. In SVM, a hyperplane is selected to best separate the points in the input variable space by their class, either class 0 or class 1. In two-dimensions, you can visualize this as a line and let's assume that all of our input points can be completely separated by this line. The SVM learning algorithm finds the coefficients that results in the best separation of the classes by the hyperplane.

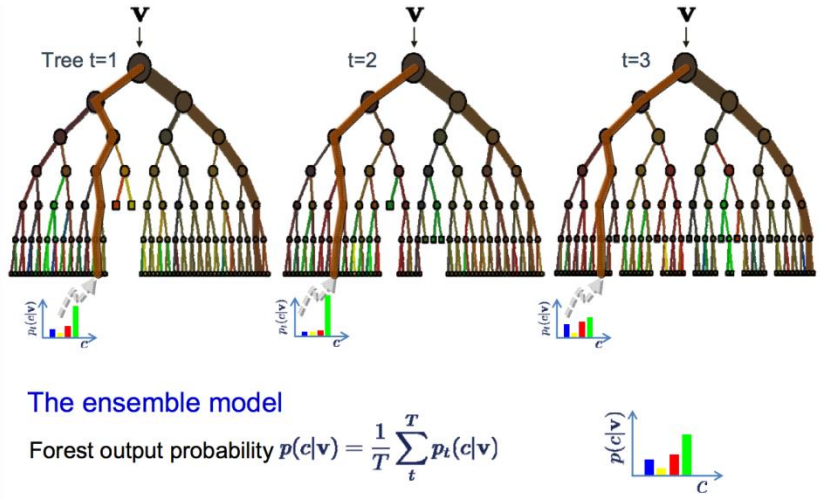


Bagging and Random Forest

Random Forest is one of the most popular and most powerful machine learning algorithms. It is a type of ensemble machine learning algorithm called Bootstrap Aggregation or bagging.

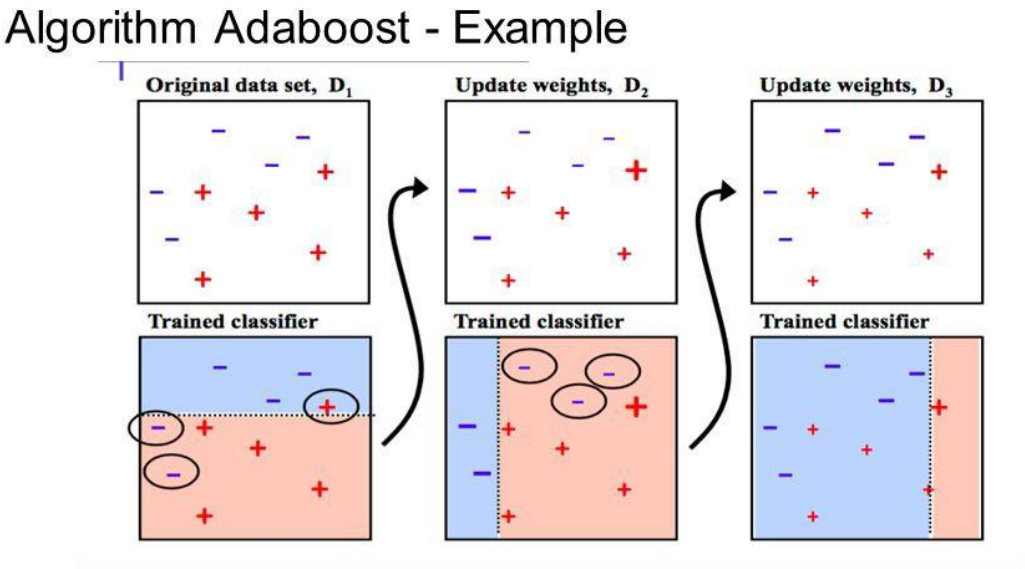
The bootstrap is a powerful statistical method for estimating a quantity from a data sample. Such as a mean. You take lots of samples of your data, calculate the mean, then average all of your mean values to give you a better estimation of the true mean value.

In bagging, the same approach is used, but instead for estimating entire statistical models, most commonly decision trees. Multiple samples of your training data are taken then models are constructed for each data sample. When you need to make a prediction for new data, each model makes a prediction and the predictions are averaged to give a better estimate of the true output value.



Boosting and AdaBoost

Boosting is an ensemble technique that attempts to create a strong classifier from a number of weak classifiers. This is done by building a model from the training data, then creating a second model that attempts to correct the errors from the first model. Models are added until the training set is predicted perfectly or a maximum number of models are added.



courtesy to Alexander Ihler http://sli.ics.uci.edu/Classes/2012F-273a?action=download&upname=10-ensembles.pdf