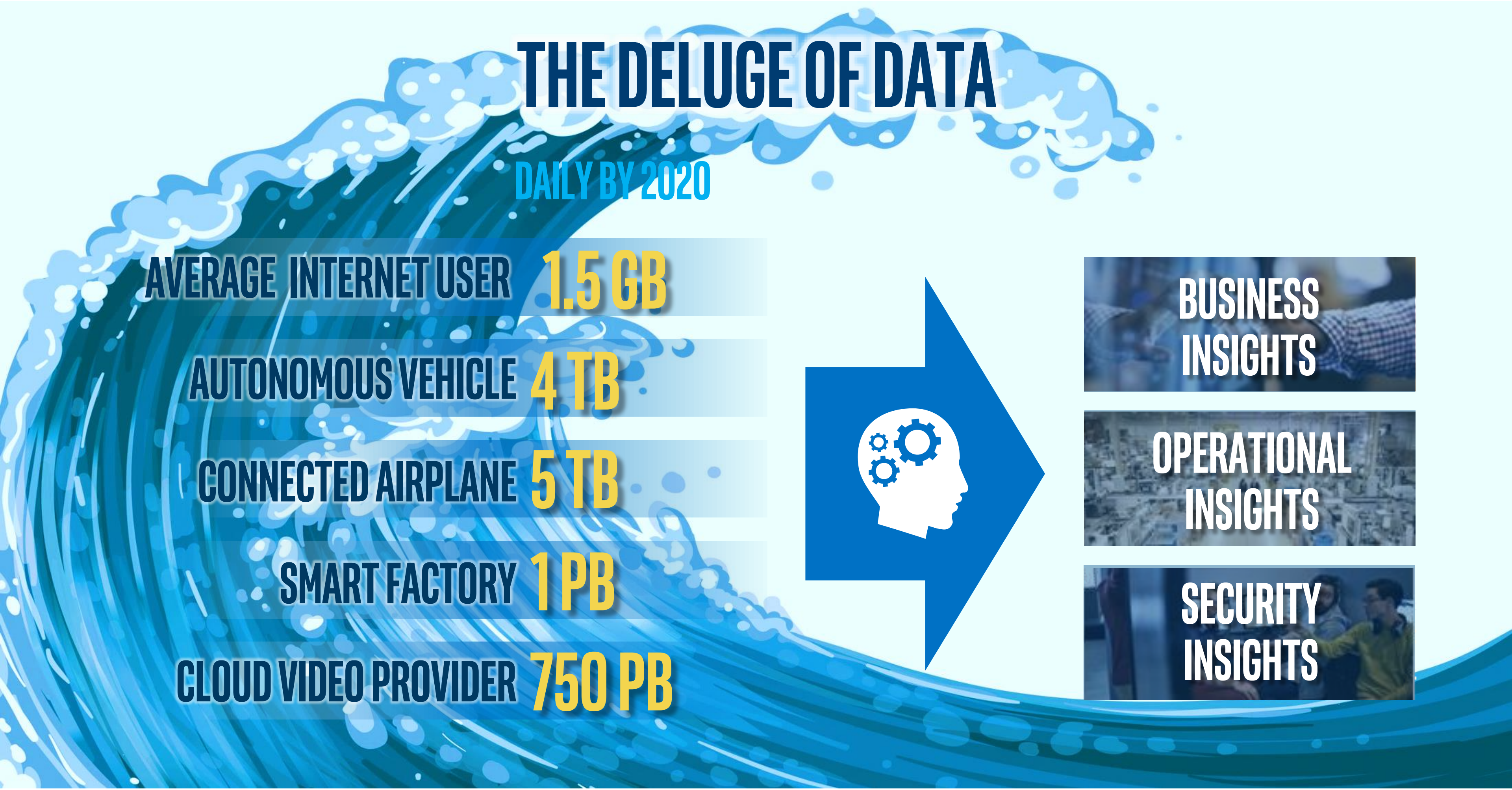
INTRODUCTION TO AI – AI FOR EVERYONE

Outline

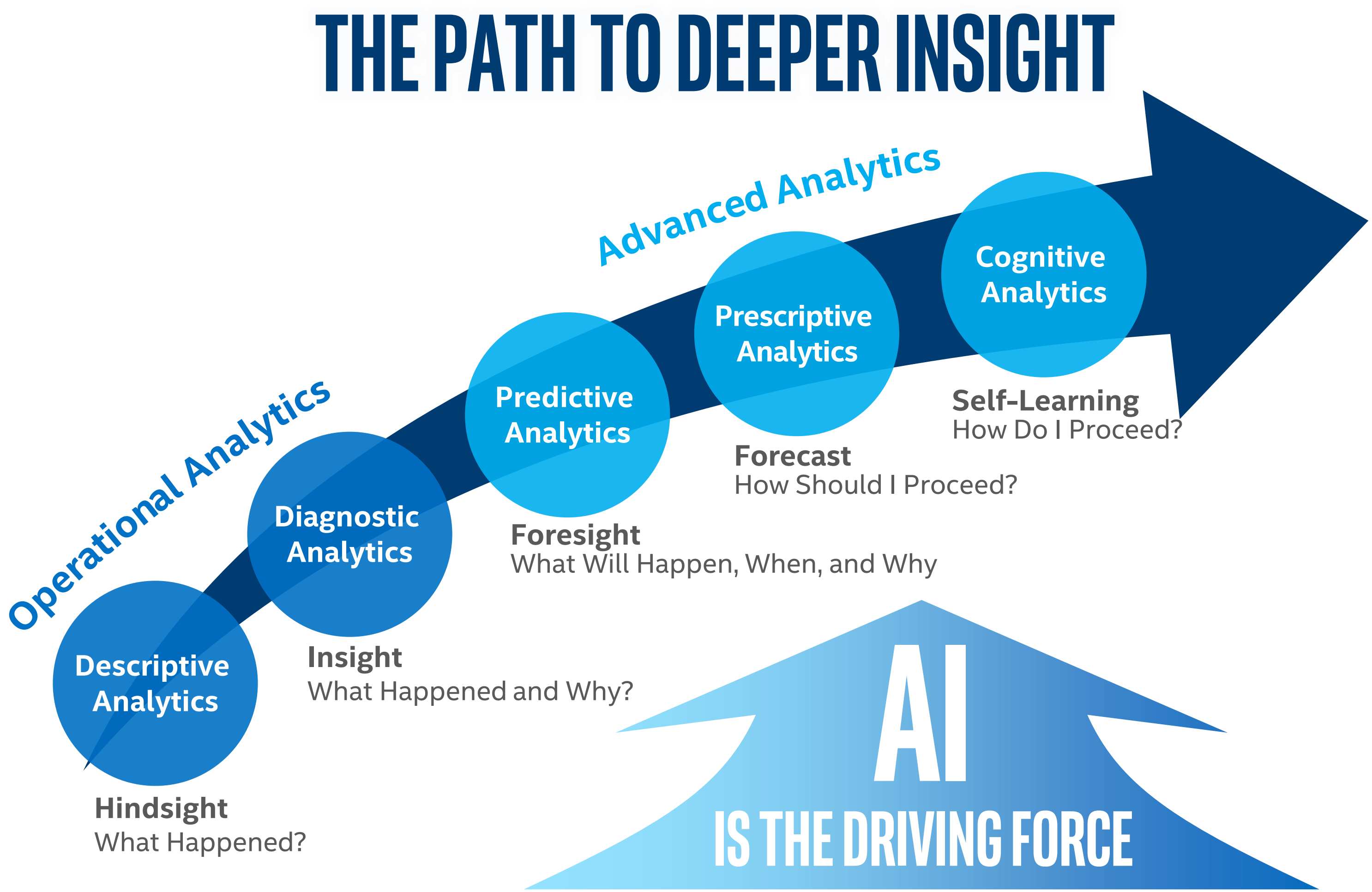
1. What is AI
2. Background of AI
3. Machine Learning
4. Deep Learning
5. Summary


THE DELUGE OF DATA
DAILY BY 2020
AVERAGE INTERNET USER 1.5 GB
AUTONOMOUS VEHICLE 4 TB
CONNECTED AIRPLANE 5 TB
SMART FACTORY 1 PB
CLOUD VIDEO PROVIDER 750 PB
BUSINESS INSIGHTS
OPERATIONAL INSIGHTS
SECURITY INSIGHTS

AI WILL TRANSFORM

CONSUMER	HEALTH	FINANCE	RETAIL	GOVERNMENT	ENERGY	TRANSPORT	INDUSTRIAL	OTHER
Smart Assistants	Enhanced Diagnostics	Algorithmic Trading	Support	Defense	Oil & Gas Exploration	In-Vehicle Experience	Factory Automation	Advertising
Chatbots	Drug Discovery	Fraud Detection	Experience	Data Insights	Smart Grid	Automated Driving	Predictive Maintenance	Education
Search	Patient Care	Research	Marketing	Safety & Security	Operational Improvement	Aerospace	Precision Agriculture	Gaming
Personalization	Research	Personal Finance	Merchandising	Resident Engagement	Conservation	Shipping	Field Automation	Professional & IT Services
Augmented Reality	Sensory Aids	Risk Mitigation	Loyalty	Smarter Cities		Search & Rescue		Telco/Media
Robots			Supply Chain					Sports
			Security					


THE PATH TO DEEPER INSIGHT
Operational Analytics
Advanced Analytics
Descriptive Analytics
Hindsight
What Happened?
Diagnostic Analytics
Insight
What Happened and Why?
Predictive Analytics
Foresight
What Will Happen, When, and Why
Prescriptive Analytics
Forecast
How Should I Proceed?
Cognitive Analytics
Self-Learning
How Do I Proceed?
AI
IS THE DRIVING FORCE

AI ADOPTION IS NASCENT

According to a recent Gartner survey...

46%

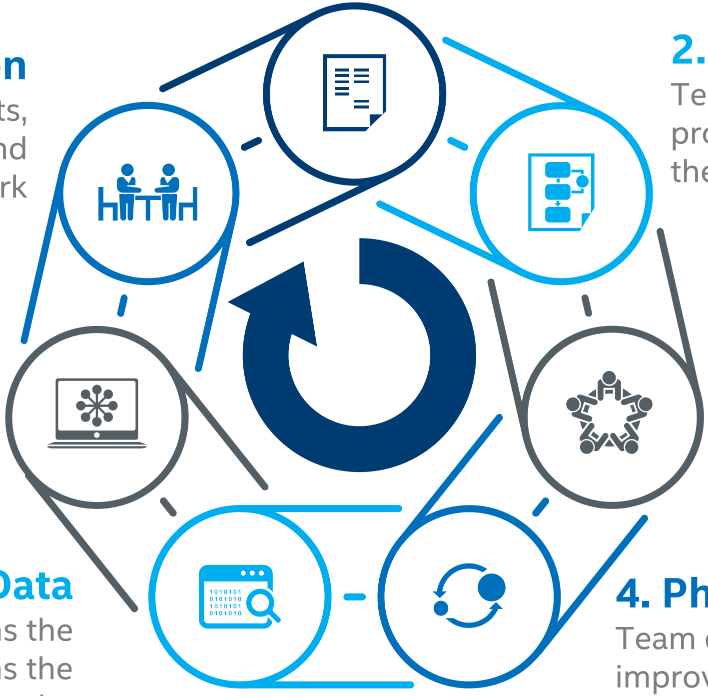
of Chief Information Officers (CIOs) have developed plans to implement AI, but **only**

4%

have implemented AI so far.

THE AI LIFECYCLE

1. Define the Challenge



2. Approach

Team breaks down the defined business problem into workable steps to translate the right data to achieve results

3. Expertise

A team of management sponsors, data scientists, data engineers, solution architects, and domain experts identifies the right data and works to translate the data to achieve results

4. Philosophy

Team embraces fail-fast continuous improvement practices to evaluate their success in translating data to achieve results

5. Source Data

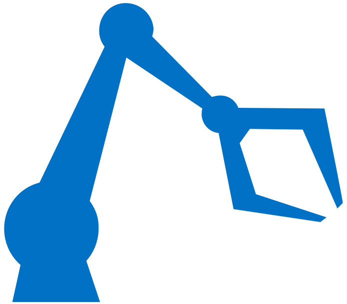
Team understands and obtains the right data that explains the business problem to achieve results

6. Infrastructure

Organization secures hardware and software infrastructure that supports data processing in a timely manner

7. Organization

Organization embraces data insights, sponsors properly resourced teams, and prioritizes analytic development work



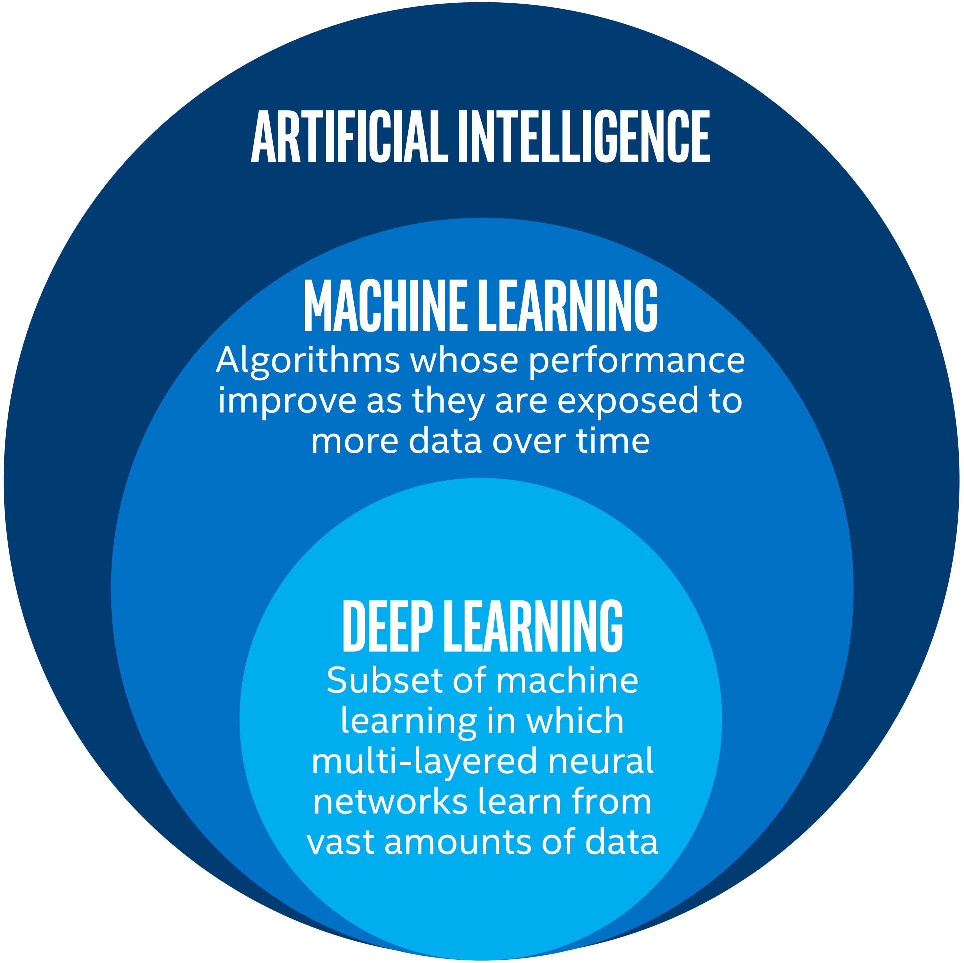
WHICH APPROACH IS RIGHT?

A large **manufacturer** uses data to improve their operations, with each challenge using a different approach to deliver maximum business value at the lowest possible cost

CHALLENGE	BEST APPROACH	APPROACH	ANSWER
How many widgets should we manufacture?	Analyze historical supply/demand	Analytics/ Business Intelligence	10,000
What will our yield be?	Algorithm that correlates many variables to yield	Statistical/ Machine Learning	At current conditions, yield will be at 90% with 10% loss expected
Which widgets have visual defects?	Algorithm that learns to identify defects in images	Deep Learning	Widget 1003, Widget 1094 . . .

ARTIFICIAL INTELLIGENCE

is the ability of machines to learn from experience, without explicit programming, in order to perform cognitive functions associated with the human mind

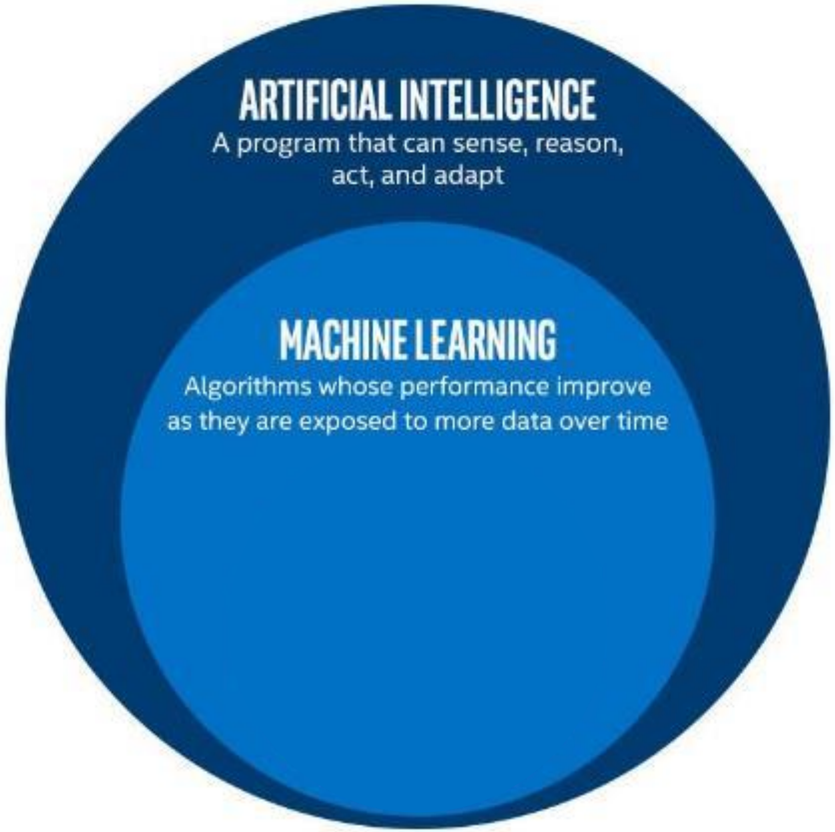


Artificial Intelligence

- “A branch of computer science dealing with the simulation of intelligent behavior in computers.” (Merriam-Webster)
- “Colloquially, the term ‘artificial intelligence’ is applied when a machine mimics ‘cognitive’ functions that humans associate with other human minds, such as ‘learning' and ‘problem solving’.” (Wikipedia)

Machine Learning

“The study and construction of programs that are *not explicitly programmed*, but learn patterns as they are exposed to more data over time.



Machine Learning

These programs learn from repeatedly seeing data, rather than being explicitly programmed by humans.

Machine Learning Terminology

This example is learning to classify a species from a set of measurement features.

Features:
Attributes of the data.

Target:
Column to be predicted.

sepal length	sepal width	petal length	petal width	species
6.7	3.0	5.2	2.3	virginica
6.4	2.8	5.6	2.1	virginica
4.6	3.4	1.4	0.3	setosa
6.9	3.1	4.9	1.5	versicolor
4.4	2.9	1.4	0.2	setosa
4.8	3.0	1.4	0.1	setosa
5.9	3.0	5.1	1.8	virginica
5.4	3.9	1.3	0.4	setosa
4.9	3.0	1.4	0.2	setosa
5.4	3.4	1.7	0.2	setosa

Two Main Types of Machine Learning

	Dataset	Goal	Example
Supervised Learning	Has a target column	Make predictions	Fraud detection
Unsupervised Learning	Does not have a target column	Find structure in the data	Customer segmentation

Machine Learning Example – fraud detection

- Suppose you wanted to identify fraudulent credit card transactions.
- You could define features to be:
  - Transaction time
  - Transaction amount
  - Transaction location
  - Category of purchase
- The algorithm could learn what feature combinations suggest unusual activity.

Machine Learning Limitations

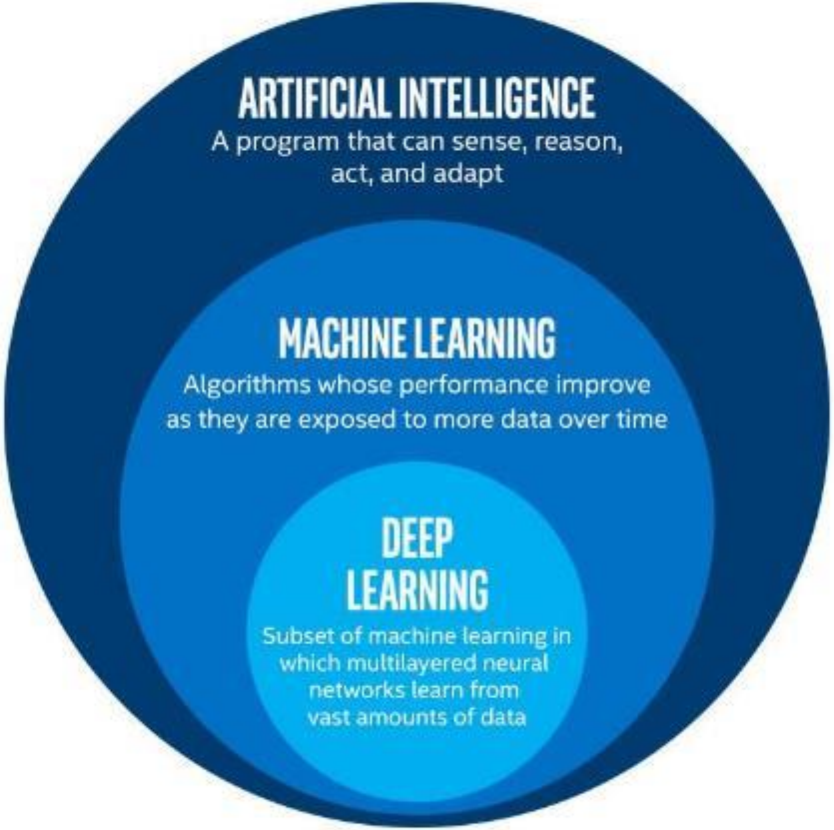
- Suppose you wanted to determine if an ima
a dog.
- What features would you use?
- This is where Deep Learning can come in.

Dog and cat recognition

Deep Learning

“Machine learning that involves using very complicated models called “deep neural networks”."

Models determine best representation of original data; in classic machine learning, humans must do this.



MACHINE VS. DEEP LEARNING

MACHINE LEARNING

How do you engineer the best features?

$N \times N$

$(f_1, f_2, \dots, f_K)$

Roundness of face
Dist between eyes
Nose width
Eye socket depth
Cheek bone structure
Jaw line length
...etc.

CLASSIFIER ALGORITHM

SVM
Random Forest
Naïve Bayes
Decision Trees
Logistic Regression
Ensemble methods

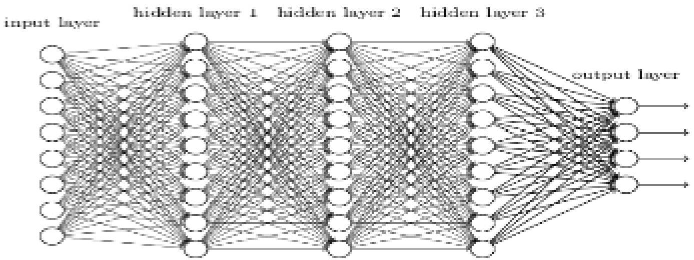
Arjun

DEEP LEARNING

How do you guide the model to find the best features?

$N \times N$

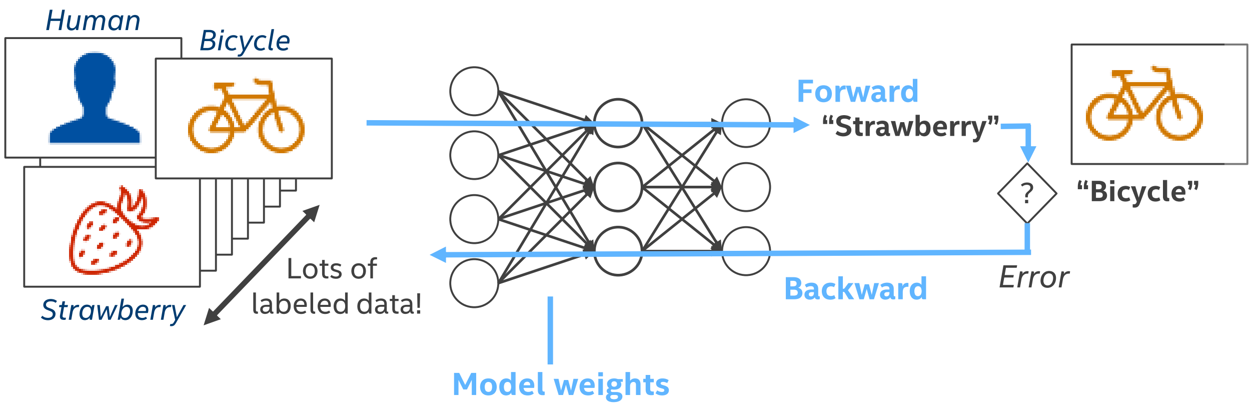
NEURAL NETWORK



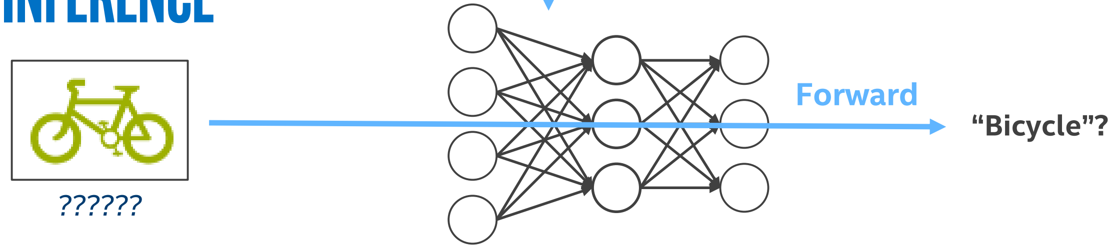
Arjun

DEEP LEARNING BASICS

TRAINING

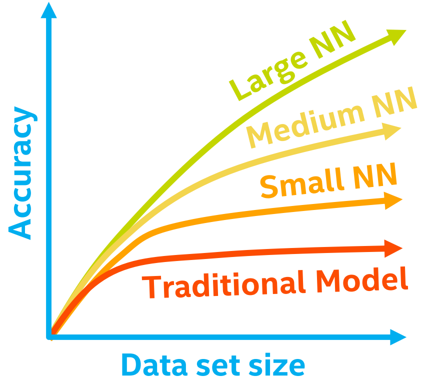


INFERENCE



DID YOU KNOW?

Training with a large data set AND deep (many layered) neural network often leads to the highest accuracy inference



Data Science Workflow

Problem Statement	What problem are you trying to solve?
Data Collection	What data do you need to solve it?
Data Exploration & Preprocessing	How should you clean your data so your model can use it?
Modeling	Build a model to solve your problem?
Validation	Did I solve the problem?
Decision Making & Deployment	Communicate to stakeholders or put into production?

MACHINE LEARNING

Machines Learn in Two Ways:

Supervised Learning & Unsupervised Learning

Supervised Learning

We train the model. We feed the model with correct answers. Model Learns and finally predicts.

We feed the model with “ground truth”.

Unsupervised Learning

Data is given to the model. Right answers are not provided to the model. The model makes sense of the data given to it.

Can teach you something you were probably not aware of in the given dataset.

Types of Supervised and Unsupervised learning

SUPERVISED

- CLASSIFICATION
- REGRESSION

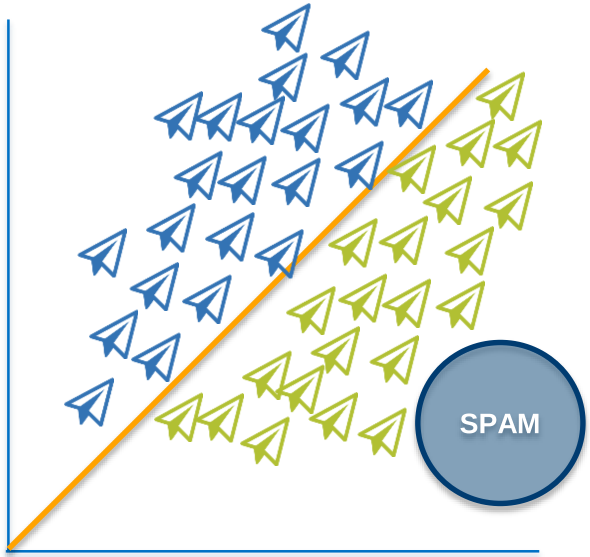
UNSUPERVISED

- CLUSTERING
- RECOMMENDATION

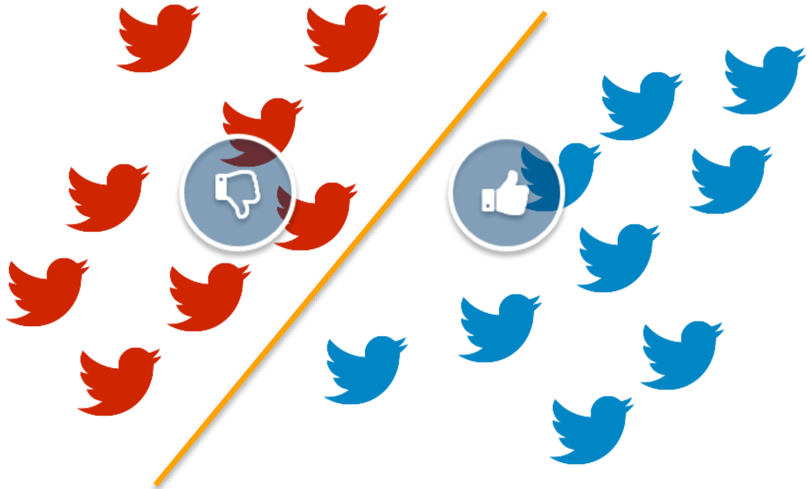
EXAMPLES OF SUPERVISED LEARNING - CLASSIFICATION

Predict a **label** for an entity with a given set of features.

PREDICTION



SENTIMENT ANALYSIS



Supervised Learning Overview

Training: Train a model with known data.

Data with answers (features and labels) + **Model** $\xrightarrow{\textbf{Fit}}$ **Trained Model**

Inference: Feed unseen data into trained model to make predictions.

Data without answers (features only) + **Trained Model** $\xrightarrow{\textbf{Predict}}$ **Predicted Answer**

Which Model?

Some considerations when choosing are:

- Time needed for training
- Speed in making predictions
- Amount of data needed
- Type of data
- Problem complexity
- Ability to solve a complex problem
- Tendency to overcomplicate a simple one

Evaluation Metric

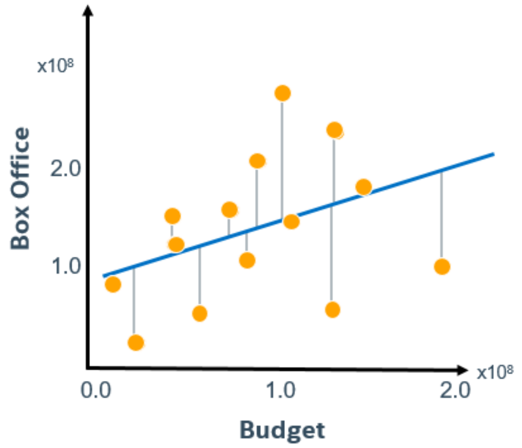
There are many metrics available* to measure performance, such as:

- **Accuracy**: how well predictions match true values.

Accuracy target

- **Mean Squared Error**: average square distance between prediction and true value.

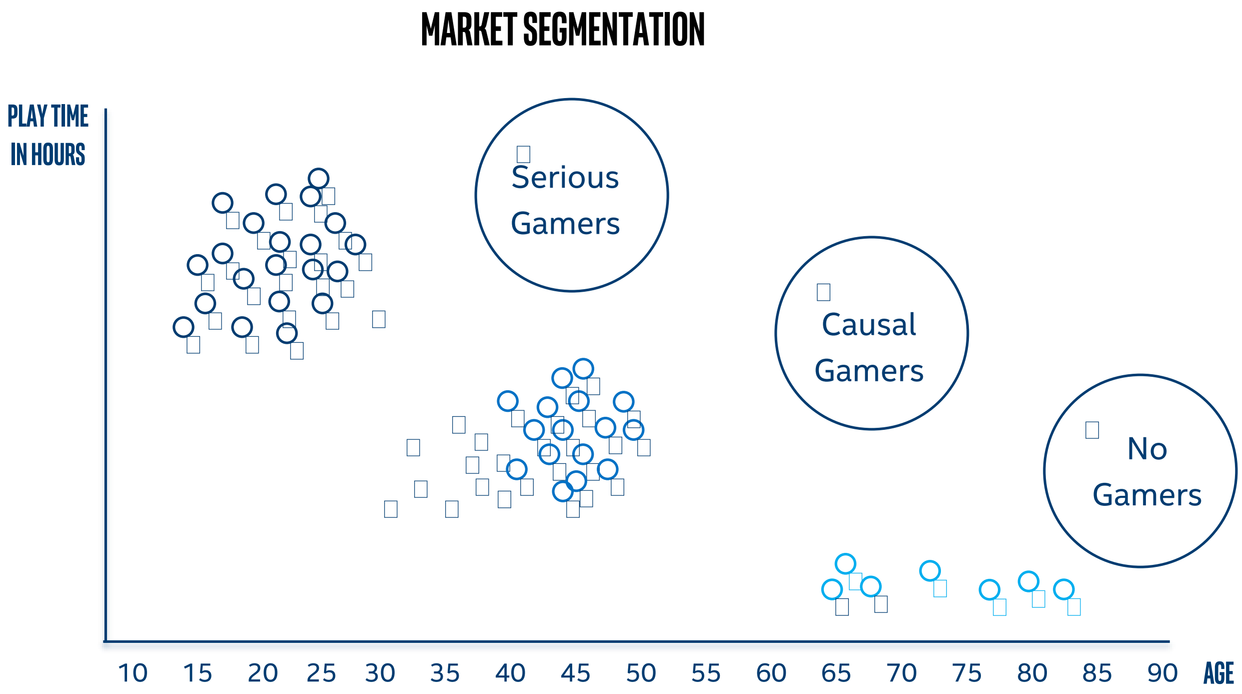
$$\min_{\beta_0,\beta_1} \frac{1}{m} \sum_{i=1}^{m} \left(\left(\beta_0 + \beta_1 x_{obs}^{(i)} \right) - y_{obs}^{(i)} \right)^2$$



**The wrong metric can be misleading or not capture the real problem.*

EXAMPLE OF UNSUPERVISED LEARNING - CLUSTERING

Group entities with similar features

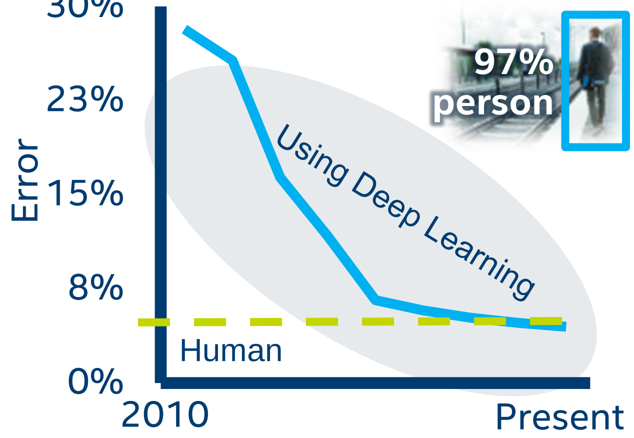


DEEP LEARNING

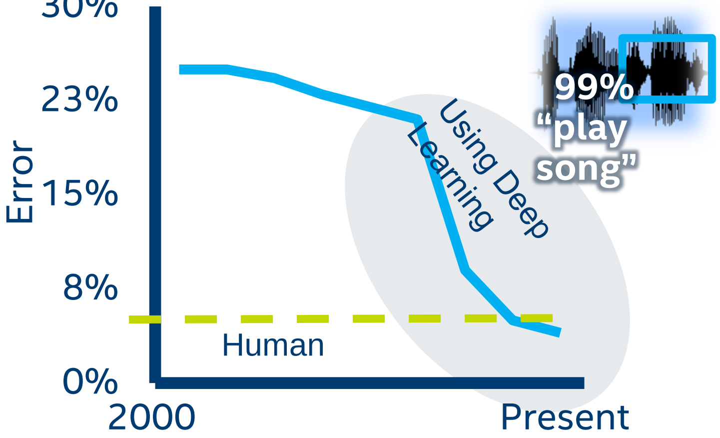
DEEP LEARNING BREAKTHROUGHS

Machines able to meet or exceed human image & speech recognition

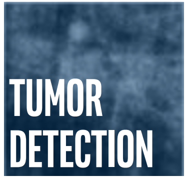
IMAGE RECOGNITION



SPEECH RECOGNITION



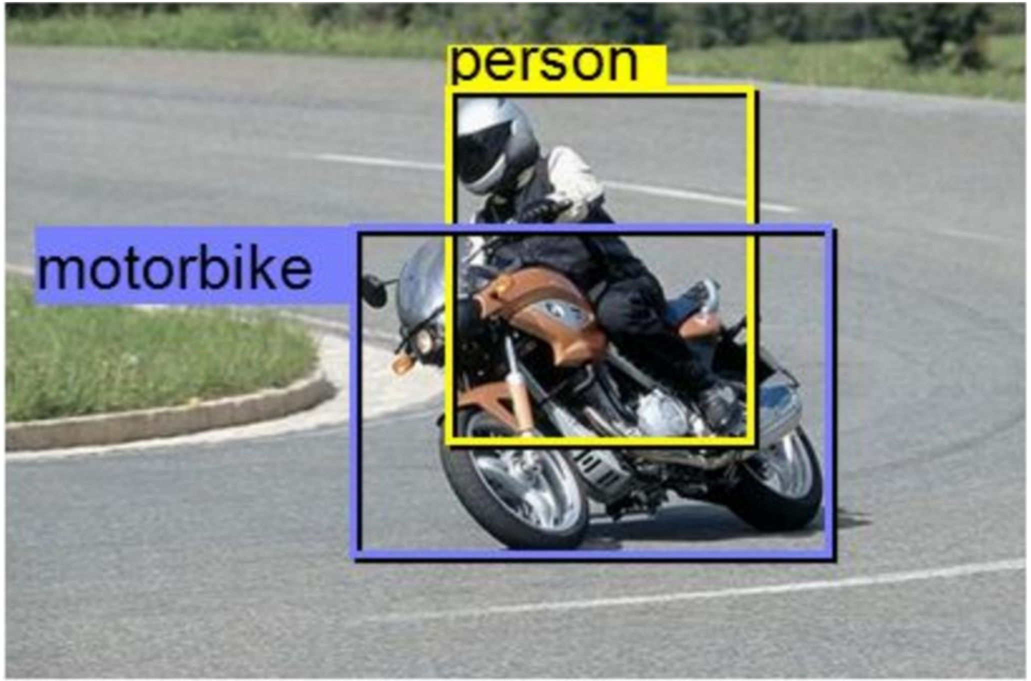
e.g.



Classification and Detection

Detect and label the image

- Person
- Motor Bike

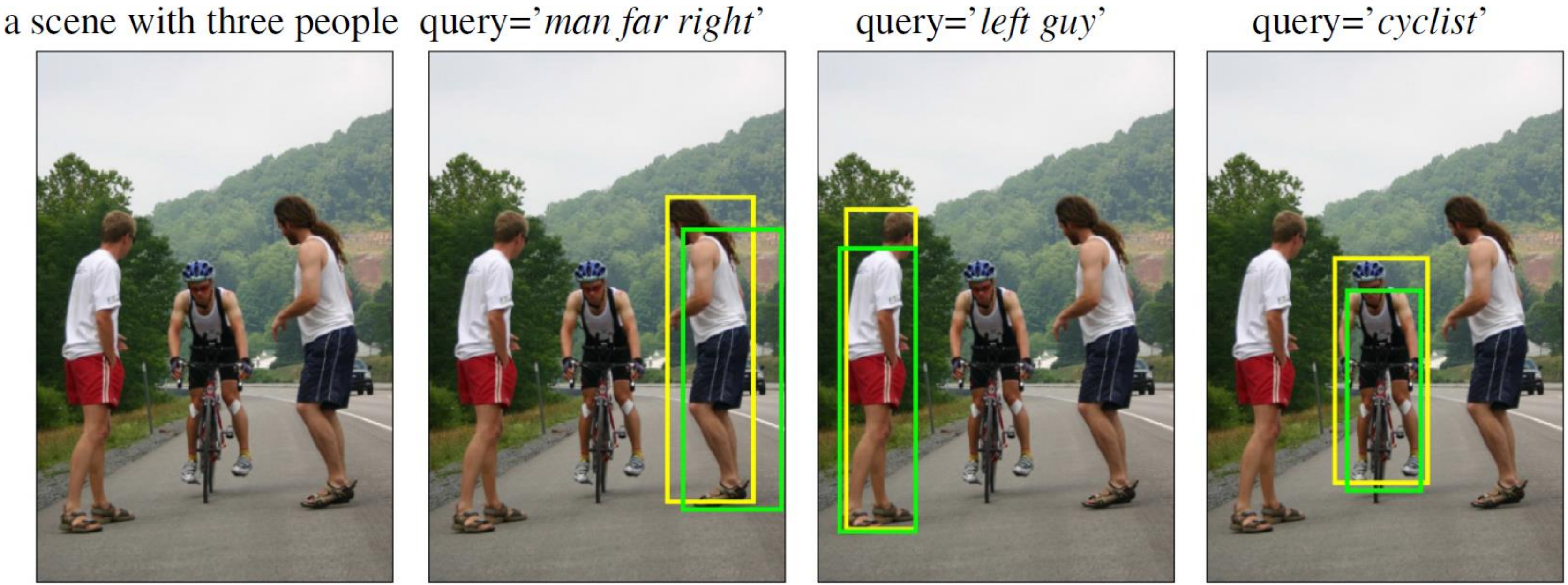


Semantic Segmentation

Label every pixel

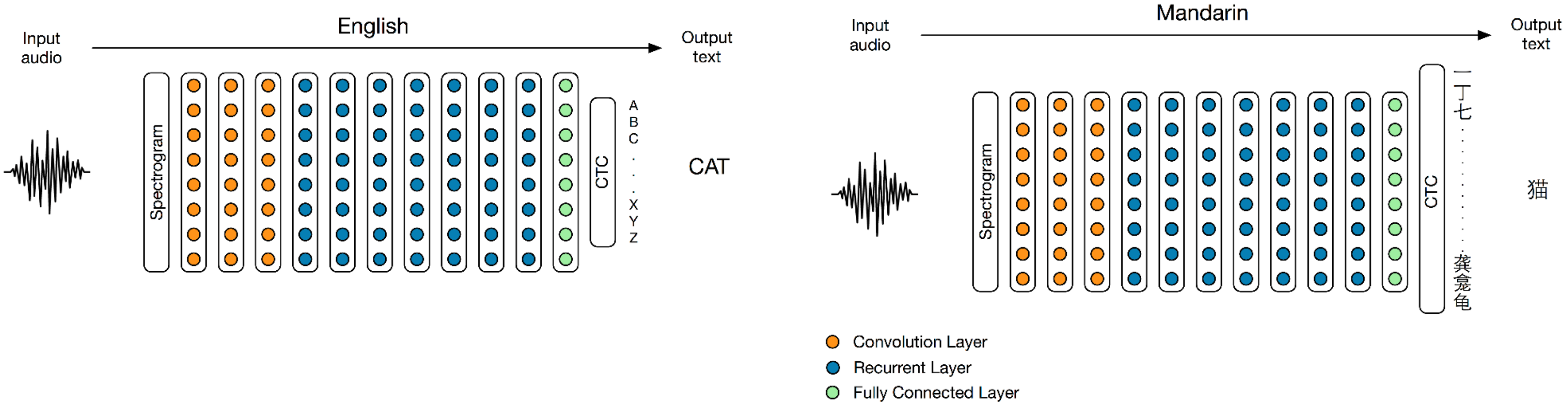
https://people.eecs.berkeley.edu/~jhoffman/talks/lsda-baylearn2014.pdf

Natural Language Object Retrieval



http://arxiv.org/pdf/1511.04164v3.pdf

Speech Recognition and Language Translation



The same architecture is used for English and Mandarin Chinese speech recognition

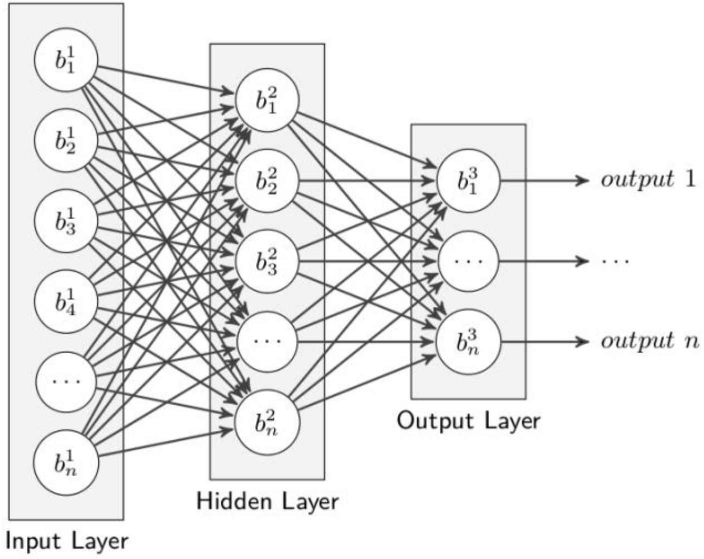
http://svail.github.io/mandarin/

NEURAL NETWORKS CONNECTIVITY

Fully Connected Network

More complicated problems can be solved by connecting multiple neurons together and using more complicated activation functions.

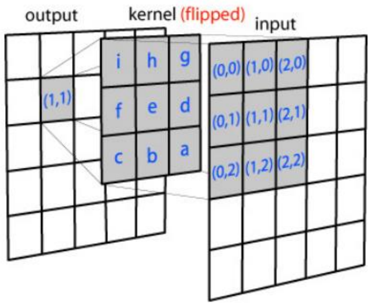
- Organized into layers of neurons.
- Each neuron is connected to every neuron in the previous layer.
- Each layer transforms the output of the previous layer and then passes it on to the next.
- Every connection has a separate weight

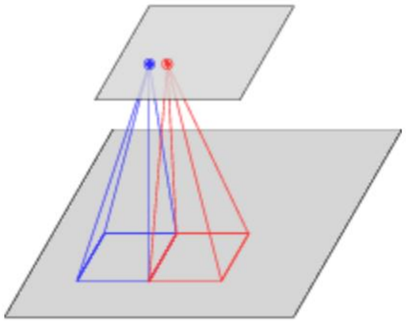


Convolutional Neural Network

Convolutional neural networks reduce the required computation and are good for detecting features.

- Each neuron is connected to a small set of nearby neurons in the previous layer
- The same set of weights are used for each neuron
- Ideal for spatial feature recognition, Ex: Image recognition
- Cheaper on resources due to fewer connections





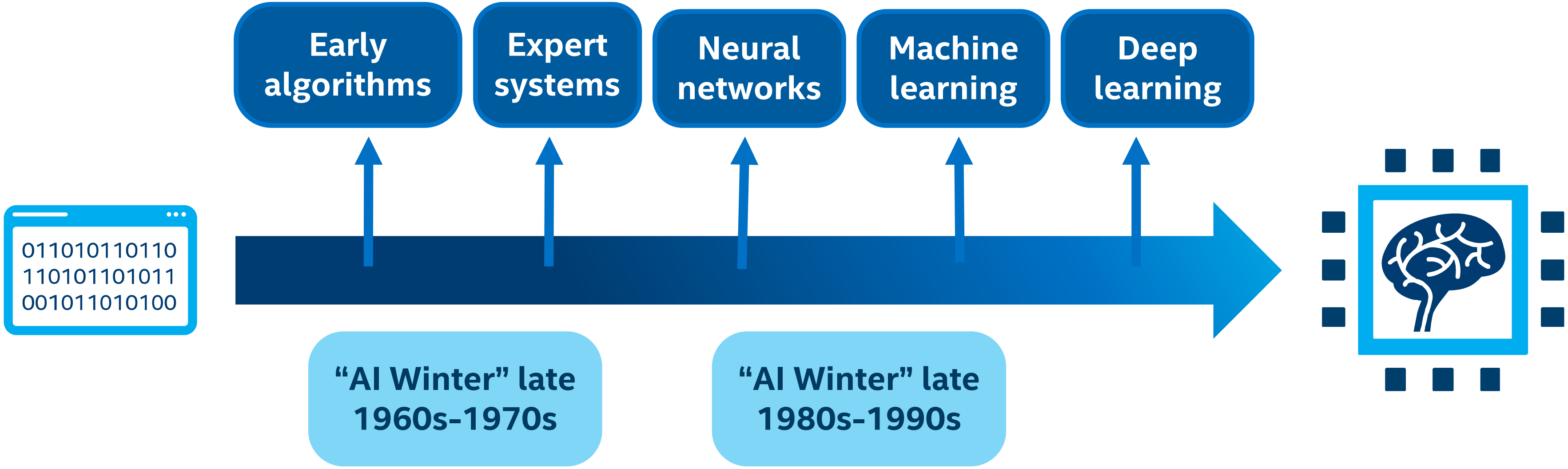
Summary

- The differences between machine learning and deep learning are:
  - Machine learning requires humans to engineer the features and the algorithms that will learn either in supervised or unsupervised modes.
  - Deep learning where the algorithms represent a variety of connection between computational nodes – the so called neural-networks.
- The differences between training and inference are:
  - Training refers to the process where data is processed by an algorithm to produce a model that is consistent with the features of the given data.
  - Inference refers to the process where a trained model when presented with new data can make predictions.

HISTORY

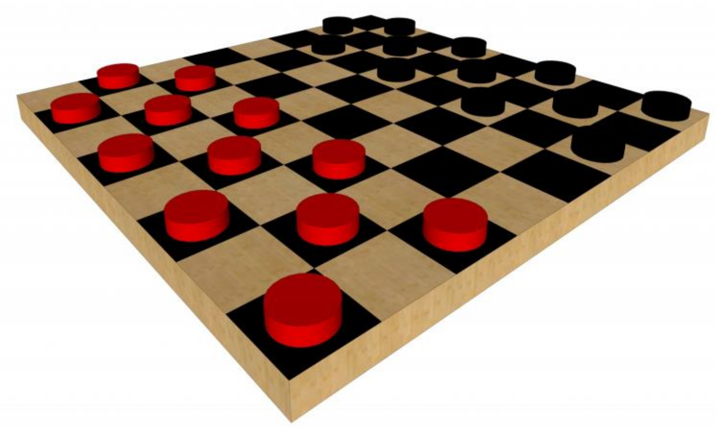
History of AI

AI has experienced several hype cycles, where it has oscillated between periods of excitement and disappointment.



1950s: Early AI

- 1950: Alan Turing developed the Turing test to test a machines ability to exhibit intelligent behavior.
- 1956: Artificial Intelligence was accepted as a field at the Dartmouth Conference.
- 1957: Frank Rosenblatt invented the perceptron algorithm. This was the precursor to modern neural networks.
- 1959: Arthur Samuel published an algorithm for a checkers program using machine learning.



The First “AI Winter”

- 1966: ALPAC committee evaluated AI techniques for machine translation and determined there was little yield from the investment.
- 1969: Marvin Minsky published a book on the limitations of the Perceptron algorithm which slowed research in neural networks.
- 1973: The Lighthill report highlights AI’s failure to live up to promises.
- The two reports led to cuts in government funding for AI research leading to the first “AI Winter.”

John R. Pierce, head of ALPAC

1980's AI Boom

- Expert Systems - systems with programmed rules designed to mimic human experts.
- Ran on mainframe computers with specialized programming languages (e.g. LISP).
- Were the first widely-used AI technology, with two-thirds of "Fortune 500" companies using them at their peak.
- 1986: The "Backpropogation" algorithm is able to train multi-layer perceptrons leading to new successes and interest in neural network research.

Early expert systems machine

Another AI Winter (late 1980's – early 1990s)

- Expert systems' progress on solving business problems slowed.
- Expert systems began to be melded into software suites of general business applications (e.g. SAP, Oracle) that could run on PCs instead of mainframes.
- Neural networks didn't scale to large problems.
- Interest in AI in business declined.

Late 1990's to early 2000's: Classical Machine Learning

- Advancements in the SVM algorithm led to it becoming the machine learning method of choice.
- AI solutions had successes in speech recognition, medical diagnosis, robotics, and many other areas.
- AI algorithms were integrated into larger systems and became useful throughout industry.
- The Deep Blue chess system beat world chess champion Garry Kasparov.
- Google search engine launched using artificial intelligence technology.

IBM supercomputer

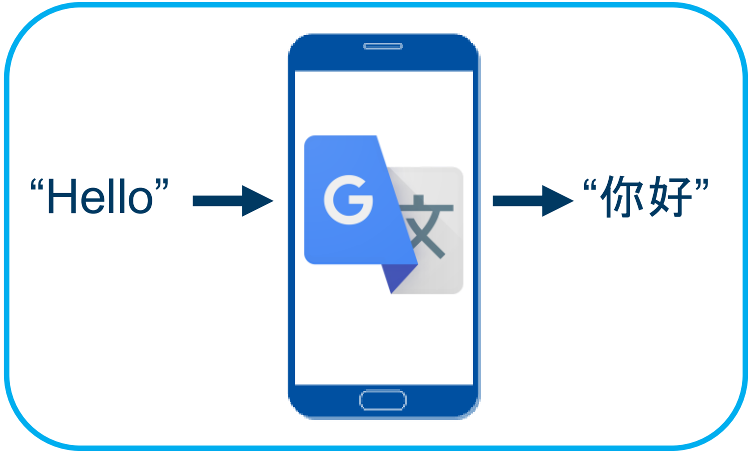
2006: Rise of Deep Learning

- 2006: Geoffrey Hinton publishes a paper on unsupervised pre-training that allowed deeper neural networks to be trained.
- Neural networks are rebranded to deep learning.
- 2009: The ImageNet database of human-tagged images is presented at the CVPR conference.
- 2010: Algorithms compete on several visual recognition tasks at the first ImageNet competition.

MODERN AI

Deep Learning Breakthroughs (2012 – Present)

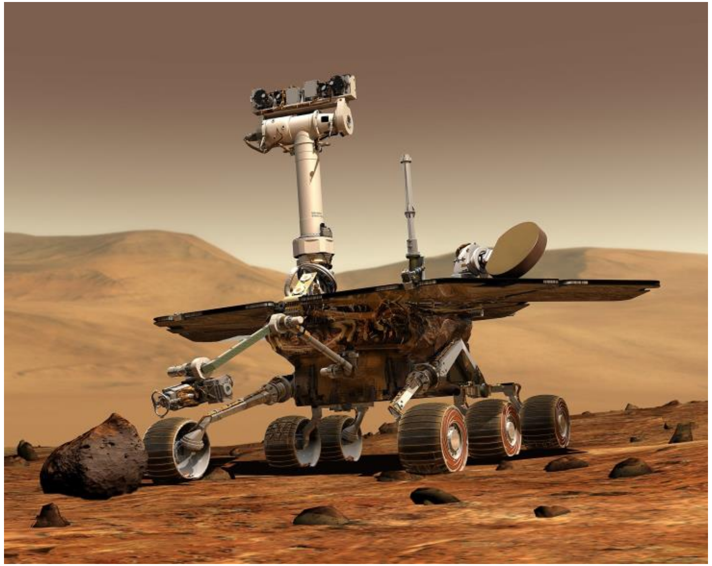
- In 2012, deep learning beats previous benchmark on the ImageNet competition.
- In 2013, deep learning is used to understand “conceptual meaning” of words.
- In 2014, similar breakthroughs appeared in language translation.
- These have led to advancements in Web Search, Document Search, Document Summarization, and Machine Translation.



Google Translate

Deep Learning Breakthroughs (2012 – Present)

- In 2014, computer vision algorithm can describe photos.
- In 2015, Deep learning platform TensorFlow* is developed.
- In 2016, DeepMind* AlphaGo, developed by Aja Huang, beats Go master Lee Se-dol.



Modern AI (2012 – Present): Deep Learning Impact

Computer vision

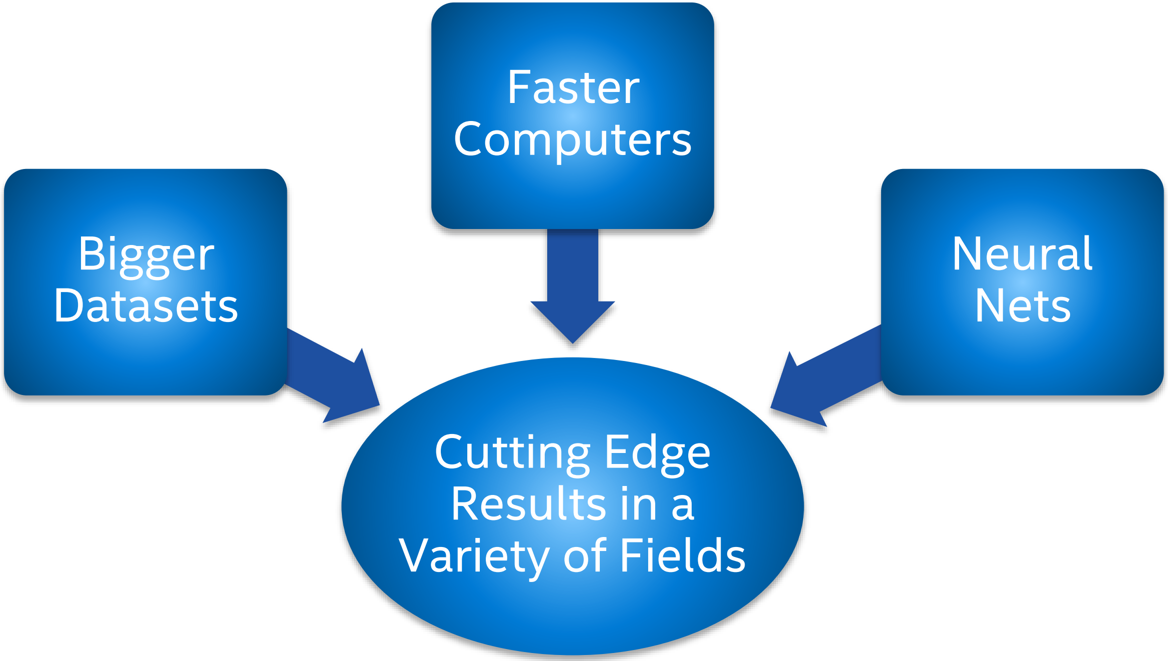
Self-driving cars:
object detection

Healthcare:
improved diagnosis

Natural language

Communication:
language translation

How Is This Era of AI Different?



Other Modern AI Factors

- Continued expansion of open source AI, especially in Python*, aiding machine learning and big data ecosystems.
- Leading deep learning libraries *open sourced*, allowing further adoption by industry.
- Open sourcing of large datasets of millions of labeled images, text datasets such as Wikipedia has also driven breakthroughs.

Transformative Changes

Health

Enhanced Diagnostics

Drug Discovery

Patient Care

Research

Sensory Aids

Industrial

Factory Automation

Predictive Maintenance

Precision Agriculture

Field Automation

Transformative Changes

Finance

Algorithmic Trading

Fraud Detection

Research

Personal Finance

Risk Mitigation

Energy

Oil & Gas Exploration

Smart Grid

Operational Improvement

Conservation

Transformative Changes

Government

Defense

Data Insights

Safety & Security

Engagement

Smarter Cities

Transport

Autonomous Cars

Automated Trucking

Aerospace

Shipping

Search & Rescue

Transformative Changes

Other

Advertising

Education

Gaming

Professional & IT Services

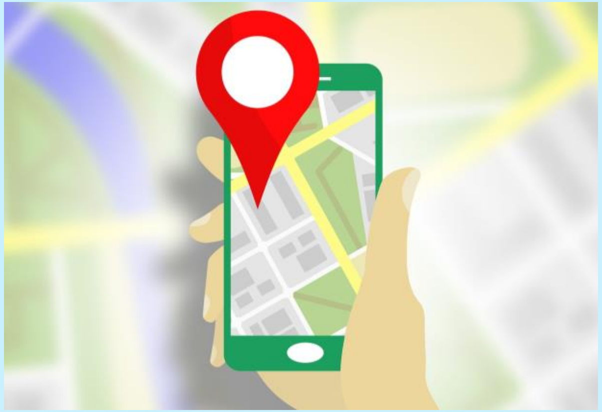
Telco/Media

Sports

APPLICATIONS

AI Omnipresence In Transportation

Navigation



Google & Waze find the fastest route, by processing traffic data.

Ride sharing

Uber & Lyft predict real-time demand using AI techniques, machine learning, deep learning.

AI Omnipresence In Social Media

Audience

Facebook & Twitter use AI to decide what content to present in their feeds to different audiences.

Content

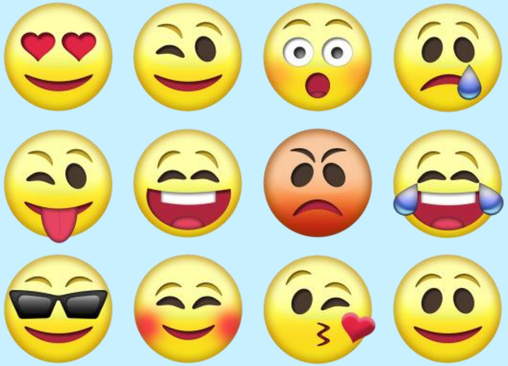


Image recognition and sentiment analysis to ensure that content of the appropriate "mood" is being served.

AI Omnipresence In Daily Life

Natural language

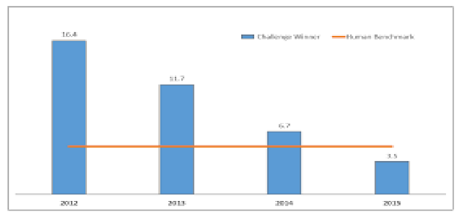
We carry around powerful natural language processing algorithms in our phones/computers.

Object detection

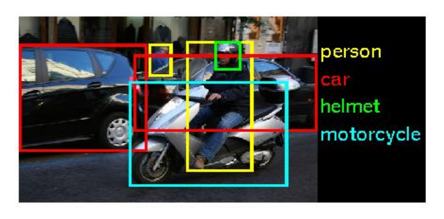
Cameras like Amazon DeepLens* or Google Clips* use object detection to determine when to take a photo.

Latest Developments: Computer Vision

Deep Learning "proven" to work for image classification.



Models outperform humans on image classification.



Object detection models beat previous benchmarks.

2012 2015 2016

Application Area: Abandoned Baggage Detection

- We can automatically detect when baggage has been left unattended, potentially saving lives.
- This system relies on the breakthroughs we discussed:
  - Cutting edge object detection.
  - Fast hardware on which to train the model

Abandoned baggage