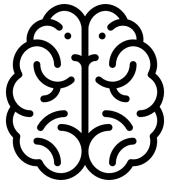
What is Machine Learning & AI?

...and why is it important?

AI: Hype or Reality?

AI is one of the most important things humanity is working on. It is more profound than electricity or fire.

Sundar Pichai, Google's CEO



Human Intelligence

- Solve problems
- Achieve goals
- Analyze & reason
- Communicate, collaborate & influence
- Consciousness,
  Emotions, Intuition, Imagination

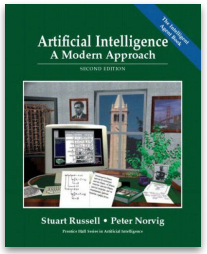
Artificial Intelligence

The ability for machines to simulate & enhance (human) intelligence

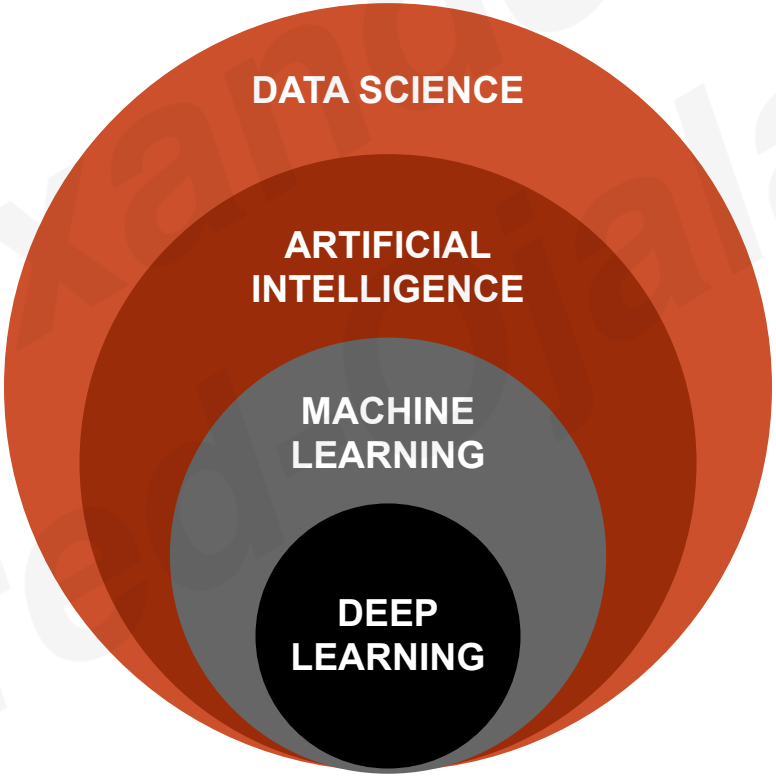
AI Definition: Academic

The designing and building of intelligent agents that receive precepts from the environment and take actions that affect that environment

Stuart Russell & Peter Norvig,
AI Professors



Machine Learning (ML) is a subfield of AI



AI is **not new**, it's been around for a loong time

Mathematical Statistics

Artificial Intelligence

Machine Learning

Deep Learning

From: AAAI Technical Report WS-98-05. Compilation copyright © 1998, AAAI (www.aaai.org). All rights reserved.

A Bayesian Approach to Filtering Junk E-Mail

Mehran Sahami* Susan Dumais† David Heckerman† Eric Horvitz†

*Gates Building 1A
Computer Science Department
Stanford University
Stanford, CA 94305-9010
sahami@cs.stanford.edu

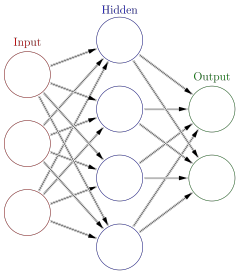
†Microsoft Research
Redmond, WA 98052-6399
{sdumais, heckerma, horvitz}@microsoft.com

Abstract

In addressing the growing problem of junk E-mail on the Internet, we examine methods for the automated construction of filters to eliminate such unwanted messages from a user's mail stream. By casting this problem in a decision theoretic framework, we are able to make use of probabilistic learning methods in conjunction with a notion of differential misclassification cost to produce filters which are especially appropriate for the nuances of this task. While this may appear, at first, to be a straight-forward text classification problem, we show that by considering domain-specific features of this problem in addition to the raw text of E-mail messages, we can produce much more accurate filters. Finally, we show the efficacy of such filters in a real world usage scenario, arguing that this technology is mature enough for deployment.

contain offensive material (such as graphic pornography), there is often a higher cost to users of actually viewing this mail than simply the time to sort out the junk. Lastly, junk mail not only wastes user time, but can also quickly fill-up file server storage space, especially at large sites with thousands of users who may all be getting duplicate copies of the same junk mail.

As a result of this growing problem, automated methods for filtering such junk from *legitimate* E-mail are becoming necessary. Indeed, many commercial products are now available which allow users to hand-craft a set of logical rules to filter junk mail. This solution, however, is problematic at best. First, systems that require users to hand-build a rule set to detect junk assume that their users are savvy enough to be able to construct robust rules. Moreover, as the nature



1700's | 1950's | 1960's | 1970's | 1980's | 1990's | 2000's | 2010's | Today | Future

1943 – The first ANN

1955 – Official term and academic recognition

1958 – Rosenblatt's Perceptron

1969 – Backpropagation

1985 – Rediscovery of Backprop

1996 – Chess victories – defeating the world champion

2012 – AlexNet wins ImageNet

2013 - Today: Deep Learning is applied almost everywhere!

What is different this time? - **4 key enablers**

1

Data availability

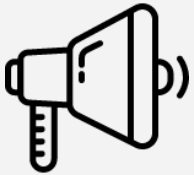
2

Computational power

3

Algorithm advancements

4



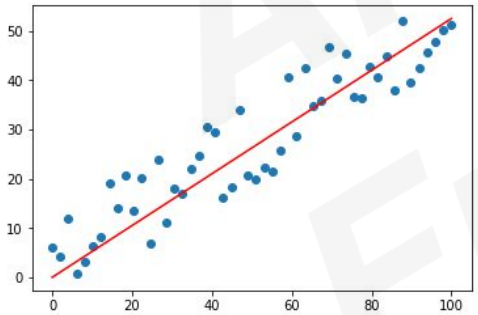
Broad public interest

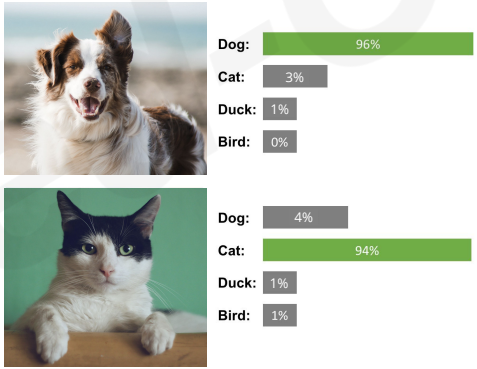
The Categories of Machine Learning

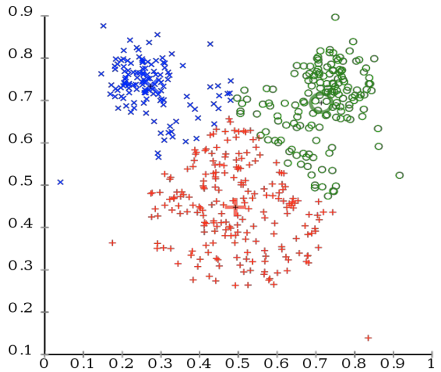
Machine Learning

- Supervised learning
  - Regression
  - Classification
- Unsupervised learning
  - Clustering, Segmentation etc

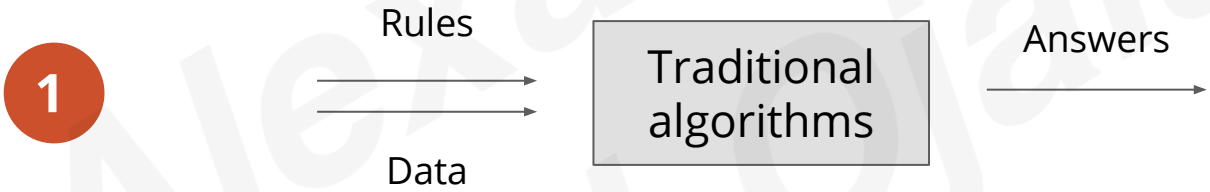
+ Reinforcement Learning







Traditional algorithms rely on rules defined by humans



Machine learning algorithms extract rules from data