Machine Learning

Introduction

"Computer Science is no more about computers than astronomy is about telescopes."

- E. W. Dijkstra

What is Computer Science?

Computer Science

Computer Science is the study of algorithms and their machine implementation.

An algorithm is a finite series of well-defined instructions for performing some task in finite time.

Is Computer Science a science?

YES!

Every time you program, you are conducting an "experiment". The hypothesis is that the task can be efficiently and correctly automated.

Theory courses in CS give you the tools for scientific investigation.

"No computer has ever been designed that is ever aware of what it's doing; but most of the time, we aren't either"

- Marvin Minsky

What is AI?

Artificial Intelligence

The AI field includes all of the following:

- Machine Learning
- Computer Vision
- Image Processing
- Pattern Recognition
- Data Mining
- Evolutionary Computation
- Knowledge Representation
- Automated Reasoning
- Natural Language Processing
- Robotics

The AI Effect

As soon as AI successfully solves a problem, the problem is no longer a part of AI.

"Practical AI successes, computational programs that actually achieved intelligent behavior, were soon assimilated into whatever application domain they were found to be useful in, and became silent partners alongside other problem-solving approaches, which left AI researchers to deal only with the "failures", the tough nuts that couldn't yet be cracked."

- Pamela McCorduck

Side Benefits of AI

- Personal Computers
- Interpreters
- Windows
- Mice
- Rapid Development Environments
- Linked List Data Type
- Automatic Storage Management
- Dynamic Programming
- Functional Programming
- Object-Oriented Programming

Defining AI

Colloquially, AI is used in at least 3 different ways:

1. The logic and actions of a computer opponent in a game
2. A machine (or computer) that mimics "cognitive" functions of a human such as perception or problem-solving
3. Business analytics and data science

Defining AI (2)

As a discipline, AI has also been defined in several ways.

Some examples from AI textbooks:

- "The art of creating machines that perform functions that require intelligence when performed by people" [Kurzweil]
- "The exciting new effort to make computers think…" [Haugeland]
- "Computational intelligence is the study of the design of intelligent agents" [Nilsson]
- "The study of mental faculties through the use of computational models" [Charniak & McDermott]

Defining AI (3)

Views of AI can be divided into four groups:

	THINK	ACT
HUMANLY	Haugeland	Charniak & McDermott
RATIONALLY	Kurzweil	Nilsson

Thinking Humanly

For a computer to think like a human, we first need to know the actual workings of the human mind.

This is **Cognitive Science**, an interdisciplinary field of psychology and computer science.

Newell and Simon's 1961 General Problem Solver (GPS) is an exemplar. Its aim was to emulate human reasoning processes computationally.

Thinking Humanly (2)

Limitations:

Humans make mistakes!

Based upon experimentation on humans to develop cognitive models. Computer Science is based upon computer experimentation.

Presupposes that the human way of reasoning and problem-solving is best.

Acting Humanly

Alan Turing (British) 1912-1954

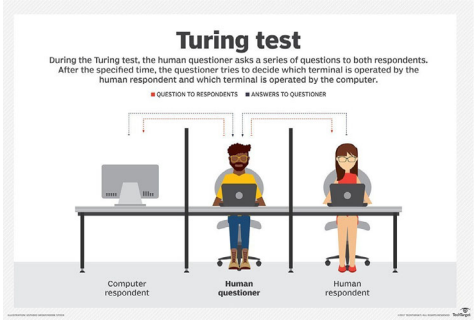
1936: *On Computable Numbers, with an application to the Entscheidungsproblem*

1939-1945: One of the British code-breakers at Bletchley Park

1950: *Computing machinery and intelligence*

1951

Acting Humanly (2)



Acting Humanly (3)

The Turing test also proposes a video signal so the questioner can test perception, and a hatch to pass physical objects.

If questioner cannot determine which is the computer respondent, then the computer passes the Turing test.

Acting Humanly (4)

To pass test, a computer would need:

- Natural Language Processing
- Knowledge Representation
- Automated Reasoning
- Machine Learning
- Computer Vision
- Robotics

These six disciplines comprise most of what is considered AI.

Acting Humanly (5)

The Turing Test is now often misapplied to computers besting humans at certain tasks. For example, "Watson passed the Turing Test for Jeopardy"…

"The good news, Dave, is that the computer's passed the Turing test. The bad news is that you've failed. "

Acting Humanly (6)

Limitations:

Restricts computer from performing better on tasks than humans (such as numeric computation).

Mimicry ≠ Intelligence.

Mimicry is not the aim of science. The Wright Brothers did not attempt to build a flying machine that would fool other birds…

Thinking Rationally

Aristotle attempted to codify rational thinking by use of **syllogisms**:

Socrates is a man
All men are mortal
Therefore, Socrates is mortal

Syllogisms were perceived as the laws by which the human mind operated.

Thinking Rationally (2)

19th century logicians developed **predicate calculus** to express statements about the relationships between objects in the world.

$$\exists x\,(\text{Barber}(x) \wedge \text{InTown}(x) \wedge \forall y\,(\text{Man}(y) \wedge \text{InTown}(y) \wedge \neg\text{Shave}(y,y) \Rightarrow \text{Shave}(x,y)))$$

By 1965, computer programs were created to solve problems stated in this notation.

Thinking Rationally (3)

Limitations:

Difficult to translate human knowledge into calculus.

Reasoning can involve uncertainty.

All possible paths of reasoning must be checked. Exhausts computational resources when solving most real-world problems.

Acting Rationally

A **rational agent** acts in such a way as to achieve the correct or best outcome.

For AI as science, this will be our approach!

- Standard of rationality is clearly defined, and does not rely upon human processes.
- Far more general than a "laws of thought" approach.

You don't understand anything until you learn it more than one way."

– Marvin Minsky

What is Learning?

What is Learning?

A simple question, but it's hard to precisely define!

Explaining the nature of learning has been at the heart of philosophy, psychology and cognitive science since their inception

Definition of Learning

"Learning is the relatively permanent change in a person's knowledge or behavior due to experience. This definition has three components: 1) the duration of the change is long-term rather than short-term; 2) the locus of the change is the content and structure of knowledge in memory or the behavior of the learner; 3) the cause of the change is the learner's experience in the environment rather than fatigue, motivation, drugs, physical condition or physiologic intervention."

- Learning in Encyclopedia of Educational Research

Definition of Learning (2)

"A *process* that leads to *change*, which occurs as a result of *experience* and increases the potential of improved performance and future learning."

- How Learning Works: Seven Research-Based Principles for Smart Teaching

"Learning is a process that occurs within nebulous environments of shifting core elements – not entirely under the control of the individual. Learning (defined as actionable knowledge) can reside outside of ourselves (within an organization or a database), is focused on connecting specialized information sets, and the connections that enable us to learn more are more important than our current state of knowing."

- Connectivism: A Learning Theory for the Digital Age

Definition of Learning? (3)

"Learning involves strengthening correct responses and weakening incorrect responses. Learning involves adding new information to your memory. Learning involves making sense of the presented material by attending to relevant information, mentally reorganizing it, and connecting it with what you already know."

- eLearning and the Science of Instruction

"A change in human disposition or capability that persists over a period of time and is not simply ascribable to processes of growth."

- The Conditions of Learning by Robert Gagne

What is Learning? (4)

A simplistic definition for us:

Learning is any process by which a system improves performance from experience.

- Herbert Simon

"Data is like garbage. You'd better know what you are going to do with it before you collect it."

- Mark Twain

What is Machine Learning?

Arthur Lee Samuel

- He's considered the Founding Father of Machine Learning
- Samuel was born in 1901 in Emporia, Kansas
- His 1959 Checkers-Playing Program was among the world's first successful self-learning programs

Machine Learning

How do we solve a problem in CS?

- As computer scientists we write a program that encodes a set of rules, but…
- It's often difficult to specify rules for real-world problems.

Is this an image of cat?

Is this an image of a dog?

Machine Learning (2)

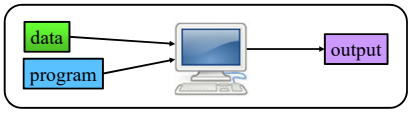
Machine Learning Systems are not "programmed" to solve a problem:

- They learn how to solve the task based on how the systems should behave
- Use trial-and-error experience to solve a problem

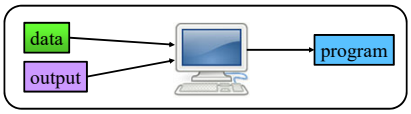
Learning means incorporating information from the learning process into the ML system.

Machine Learning (3)

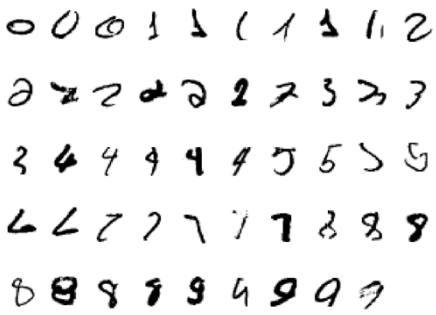
Traditional CS:



Machine Learning:



Example: What is a 2?



Machine Learning Definition

Machine Learning is the study of algorithms that
- improve their performance **P**
- at some task **T**
- with experience **E**

A well-defined learning task is given by <**P**, **T**, **E**>.

- Tom Mitchell

Machine Learning Definition (2)

Specifying the learning system requires us to Define exactly:

- what knowledge should be learned
- how this knowledge will be represented
- how the knowledge will be acquired

Desired Knowledge

The desired knowledge can be represented as a target valuation function:

$$\mathbf{V}: \mathbf{I} \rightarrow \mathbf{D}$$

It takes in information about the problem and gives back a desired decision.

Often, it is unrealistic to expect to learn the ideal function **V**, and an approximation is used instead:

$$\mathbf{V'}: \mathbf{I} \rightarrow \mathbf{D}$$

Knowledge Representation

The function **V** or **V'** must be represented symbolically in a language **L**.

L can be anything including:

- programming language
- Boolean expressions
- Arithmetic functions
- Grammar
- Etc…

Learning Methodology

We must specify an algorithm **A** which defines the way in which the system is to search the language **L** for **V** (or an acceptable **V'**).

This means we must specify a **search algorithm**.

The **search space** is all possible constructs in L.

Learning means finding **V** or **V'** in the search space.

Examples of Learning Tasks

T: Playing checkers
P: Percentage of games won
E: Playing practice games against itself

T: Recognizing hand-written words
P: Percentage of words correctly classified
E: Database of human-labeled images of handwritten words

T: Categorize email messages as spam or legitimate
P: Percentage of email messages correctly classified
E: Database of emails, some with human-given labels

Bias

Computational learning theory tell us that learning without some presuppositions is infeasible.

To learn effectively, we must limit the class of **V**'s.

Two approaches are used in machine learning:

- Language Bias
- Search Bias

Language and search bias are not mutually exclusive.
Most learning systems feature both (Combined Bias).

No Free Lunch

Wolpert and MacReady's **No Free Lunch Theorem** states that, averaged over all problems, all biases are equally good or bad.

No machine learning system can be universal.

In most systems, the bias is not explicit.

Variance

Variance is error due to using too much complexity in the learning algorithm.

Variance leads to the algorithm being highly sensitive to high degrees of variation in your training data, which can lead your model to overfit the data.

Bias vs. Variance

If you make the model more complex and add more variables, you'll lose bias but gain some variance.

If you make the model less complex with fewer variables, you lower variance but gain bias.

You don't want either high variance or high bias!

Types of Learning

- Supervised Learning
- Unsupervised Learning
- Semi-Supervised Learning
- Reinforcement Learning

Types of Learning (2)

In **supervised learning**, the system is fed carefully selected examples and counterexamples. The examples serve as a "teacher".

In **unsupervised learning**, the system builds an internal representation of the input, capturing regularities and structure in data.

Types of Learning (3)

In **semi-supervised learning,** both supervised and unsupervised learning are used.

In **reinforcement learning**, the system takes suitable actions to maximize a payoff function.

Machine Learning Paradigms

Rote learning
Hand-encoded mapping from inputs to stored representation.

Interactive learning
Human/system interaction producing explicit mapping.

Induction
Using specific examples to reach general conclusions.

Analogy
Determining correspondence between two different representations. Case-based reasoning.

Machine Learning Paradigms (2)

Clustering
Unsupervised identification of natural groups in data

Discovery
Unsupervised, specific goal not given.

Evolutionary Algorithms
Search techniques based upon processes in nature.

Classification
Predict the correct category based on observations

Regression/Prediction
Predict target value from observations

Some Terminology

Features are the number of distinct traits that can be used to describe each item in a quantitative manner.

A **sample** is an item to process or classify. It can be whatever you can describe with a fixed set of quantitative traits.

A **feature vector** is an n-dimensional vector of numerical features that represent some object.

Some Terminology (2)

Feature extraction is the preparation of the feature vector. It transforms the data in the high-dimensional space to a space of fewer dimensions.

A **training set** is the collection of data used to discover potentially predictive relationships.

Terminology Example

Problem: Classification of iris flowers

Classes: Iris Setosa,
Iris Versicolour
Iris Virginica

Dataset: 150 instances, 50 of each class

Terminology Example (2)

Instance: sepal length in cm
sepal width in cm
petal length in cm
petal width in cm

Missing values: none

Some Aspects of Learning

- Induction vs. Deduction
- Stochastic vs. Deterministic
- Clean Data vs. Noisy Data
- Discrete vs. Continuous Variables
- Symbolic vs. Non-symbolic

Induction vs. Deduction

Induction:

Specific instances ⇨ general rule

Deduction:

General rules ⇨ specific instance

Stochastic vs. Deterministic

If a model is **deterministic**, you have all of the data necessary to predict the outcome with 100% certainty.

A **stochastic** model represents a situation where uncertainty is present.

Stochastic ≠ Random

Clean vs. Noisy Data

Clean data are entirely accurate.

Noisy data *may* contain errors.

Noise has two main sources:

- implicit errors introduced by measurement tools
- random errors introduced by batch processes or experts when the data are gathered

Discrete vs. Continuous

A discrete variable is countable in a finite amount of time.

A **continuous variable** is a variable whose value is one of an uncountable set of values. The variable is a measurement or estimate of the actual value.

Symbolic vs. Non-symbolic

Symbolic learning systems use symbol manipulation and are able to manipulate complex data structures and show the steps used to learn a task.

Non-symbolic learning systems perform calculations according to principles that have been demonstrated to solve problems, without exactly understanding how to arrive at the solution.