Machine Learning

Introduction to the Course

Avisek Gupta
Postdoctoral Fellow, IAI, TCG CREST
avisek003@gmail.com

August 12, 2022

Why study Machine Learning?

- *Understanding the world around us:*
  - Deterministic Algorithms can describe Deterministic systems (Newton's laws, shortest path between two places).
  - How can we describe systems that are not Deterministic? (Recognizing human faces, navigating the real world)

- Machine Learning: The study of algorithms that search for a function which approximately models a system that is not deterministic.

"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*."

- Tom M. Mitchell, Machine Learning, McGraw Hill, 1997

One of the first uses of the term *Machine Learning*.

- *Self-teaching computers* was also alternately used.

Arthur Lee Samuel

Published in 1959

A. L. Samuel

Some Studies in Machine Learning Using the Game of Checkers

Abstract: Two machine-learning procedures have been investigated in some detail using the game of checkers. Enough work has been done to verify the fact that a computer can be programmed so that it will learn to play a better game of checkers than can be played by the person who wrote the program. Furthermore, it can learn to do this in a remarkably short period of time (8 or 10 hours of machine-playing time) when given only the rules of the game, a sense of direction, and a redundant and incomplete list of parameters which are thought to have something to do with the game, but whose correct signs and relative weights are unknown and unspecified. The principles of machine learning verified by these experiments are, of course, applicable to many other situations.

Introduction

The studies reported here have been concerned with the programming of a digital computer to behave in a way which, if done by human beings or animals, would be described as involving the process of learning. While this is not the place to dwell on the importance of machine-learning procedures, or to discourse on the philosophical aspects,[1] there is obviously a very large amount of work, now done by people, which is quite trivial in its demands on the intellect but does, nevertheless, involve some learning. We have at our command computers with adequate data-handling ability and with sufficient computational speed to make use of machine-learning techniques, but our knowledge of the basic principles of these techniques is still rudimentary. Lacking such knowledge, it is necessary to specify methods of problem solution in minute and exact detail, a time-consuming and costly procedure. Programming computers to learn from experience should eventually eliminate the need for much of this detailed programming effort.

method should lead to the development of general-purpose learning machines. A comparison between the size of the switching nets that can be reasonably constructed or simulated at the present time and the size of the neural nets used by animals, suggests that we have a long way to go before we obtain practical devices.[2] The second procedure requires reprogramming for each new application, but it is capable of realization at the present time. The experiments to be described here were based on this second approach.

- *Choice of problem*

For some years the writer has devoted his spare time to the subject of machine learning and has concentrated on the development of learning procedures as applied to games.[3] A game provides a convenient vehicle for such study as contrasted with a problem taken from life, since many of the complications of detail are removed. Checkers, rather than chess,[4–7] was chosen because the

Recent successes of Machine Learning

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature > articles > article

Article | Open Access | Published: 01 December 2021

Advancing mathematics by guiding human intuition with AI

Alex Davies, Petar Veličković, Lars Buesing, Sam Blackwell, Daniel Zheng, Nenad Tomašev, Richard Tanburn, Peter Battaglia, Charles Blundell, András Juhász, Marc Lackenby, Geordie Williamson, Demis Hassabis & Pushmeet Kohli

Nature **600**, 70–74 (2021) | Cite this article

182k Accesses | **35** Citations | **1629** Altmetric | Metrics

Abstract

The practice of mathematics involves discovering patterns and using these to formulate and prove conjectures, resulting in theorems. Since the 1960s, mathematicians have used computers to assist in the discovery of patterns and formulation of conjectures[1], most famously in the Birch and Swinnerton-Dyer conjecture[2], a Millennium Prize Problem[3]. Here

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature > articles > article

Article | Open Access | Published: 15 July 2021

Highly accurate protein structure prediction with AlphaFold

John Jumper, Richard Evans, ... Demis Hassabis + Show authors

Nature **596**, 583–589 (2021) | Cite this article

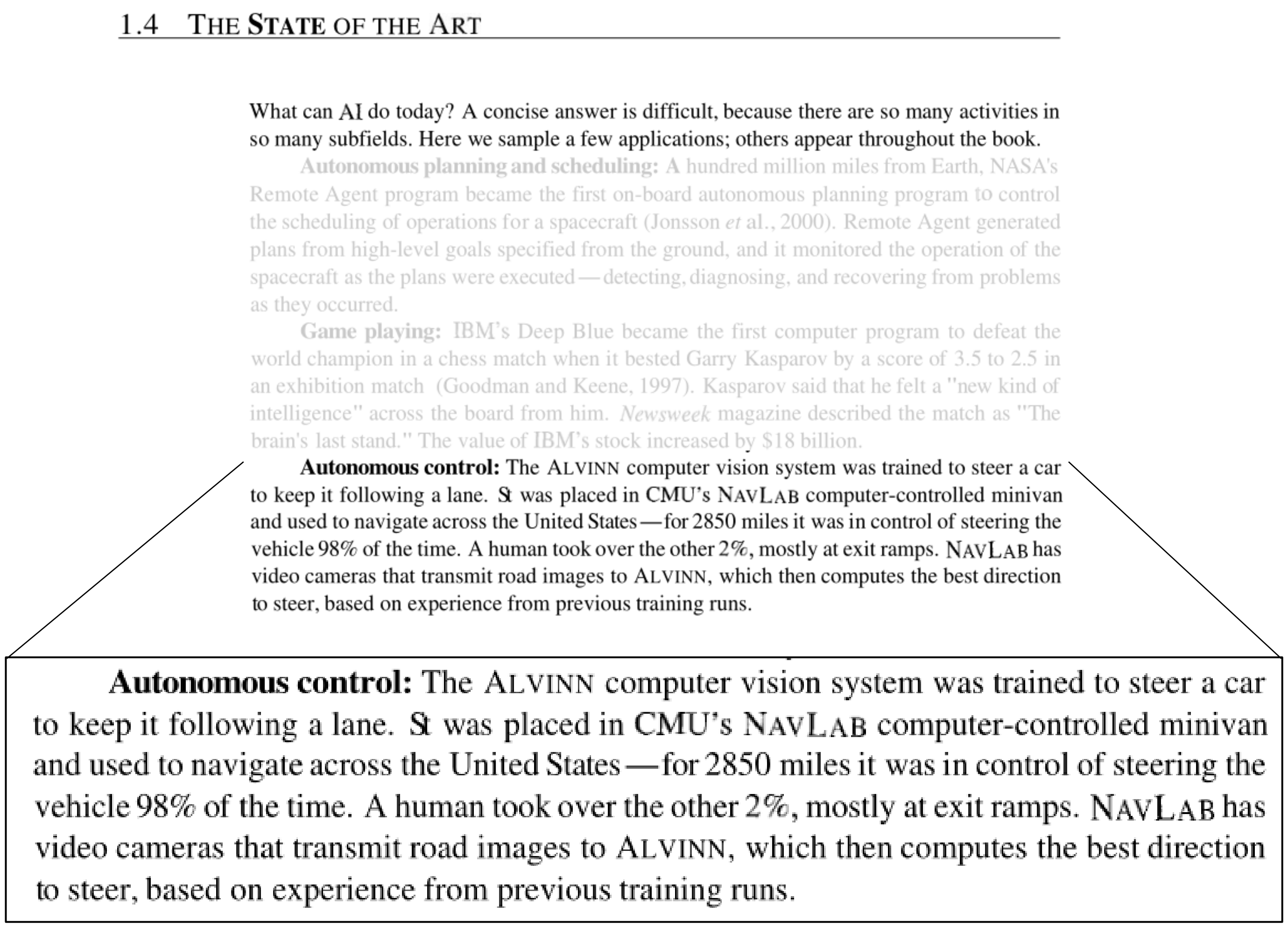
749k Accesses | **2801** Citations | **3148** Altmetric | Metrics

Abstract

Proteins are essential to life, and understanding their structure can facilitate a mechanistic understanding of their function. Through an enormous experimental effort[1,2,3,4], the structures of around 100,000 unique proteins have been determined[5], but this represents a small fraction of the billions of known protein sequences[6,7]. Structural coverage is bottlenecked by the months to years of painstaking effort required to determine a single protein structure. Accurate computational approaches are needed to address this gap and to

Rapid Advancements in Machine Learning

- From Russel S.J., Norvig P., *Artificial Intelligence, A Modern Approach (2003):*



1.4 THE STATE OF THE ART

What can AI do today? A concise answer is difficult, because there are so many activities in so many subfields. Here we sample a few applications; others appear throughout the book.

Autonomous planning and scheduling: A hundred million miles from Earth, NASA's Remote Agent program became the first on-board autonomous planning program to control the scheduling of operations for a spacecraft (Jonsson *et* al., 2000). Remote Agent generated plans from high-level goals specified from the ground, and it monitored the operation of the spacecraft as the plans were executed—detecting, diagnosing, and recovering from problems as they occurred.

Game playing: IBM's Deep Blue became the first computer program to defeat the world champion in a chess match when it bested Garry Kasparov by a score of 3.5 to 2.5 in an exhibition match (Goodman and Keene, 1997). Kasparov said that he felt a "new kind of intelligence" across the board from him. *Newsweek* magazine described the match as "The brain's last stand." The value of IBM's stock increased by $18 billion.

Autonomous control: The ALVINN computer vision system was trained to steer a car to keep it following a lane. It was placed in CMU's NAVLAB computer-controlled minivan and used to navigate across the United States—for 2850 miles it was in control of steering the vehicle 98% of the time. A human took over the other 2%, mostly at exit ramps. NAVLAB has video cameras that transmit road images to ALVINN, which then computes the best direction to steer, based on experience from previous training runs.

Autonomous control: The ALVINN computer vision system was trained to steer a car to keep it following a lane. It was placed in CMU's NAVLAB computer-controlled minivan and used to navigate across the United States—for 2850 miles it was in control of steering the vehicle 98% of the time. A human took over the other 2%, mostly at exit ramps. NAVLAB has video cameras that transmit road images to ALVINN, which then computes the best direction to steer, based on experience from previous training runs.

Researchers in 2019

Video Source: https://www.youtube.com/watch?v=IIReDnbLQAE
Courtesy: https://github.com/lzccccc

Course Outline

Syllabus:

- Regression: Linear, Non-linear, Non-parametric
- Classification: Logistic Regression, Bayes Classifier, k-NN, Support Vector Machines, Multi-Layered Perceptrons
- Tree-based Learning: Decision Trees, Random Forests, XGBoost
- Dimension Reduction: Linear and Non-linear Subspace Projections
- Clustering: k-Means, Hierarchical, Density-Based, Mean Shift, Subspace and Spectral Clustering

- Semi-Supervised Learning
- Reinforcement Learning
- Time-Series Data and Continuous Learning (3 hours)
- Interpretable and Explainable AI
- Ethics in AI

Evaluation:

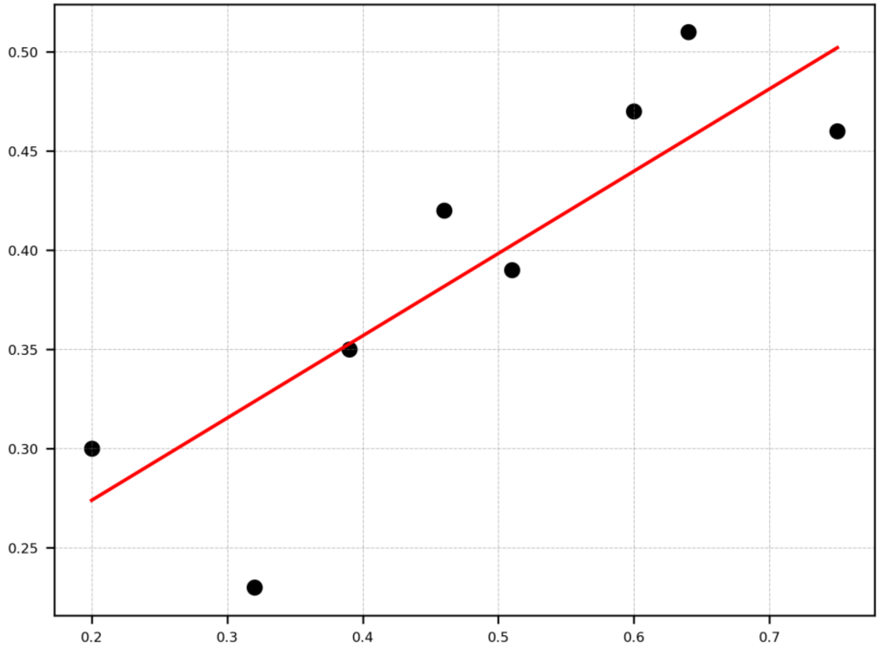
- Mid-semester (40%) and Final Examination (60%)

Initial focus: Regression and Classification

Regression: *Linear Regression*

Given data $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$,
find a ***best fit*** line:

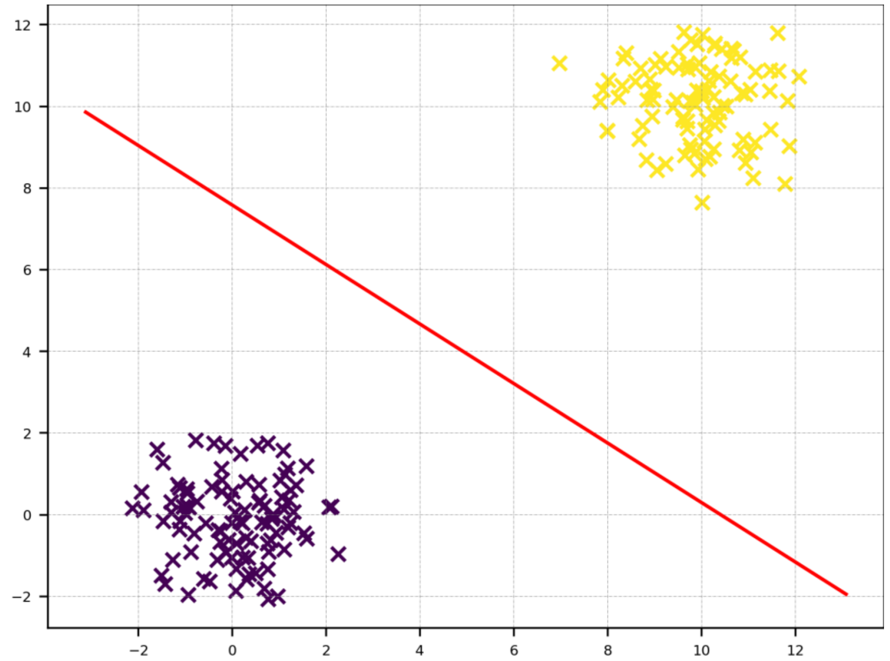
$$y = w_0 + w_1 x$$



Classification: *Logistic Regression*

Given data $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$,
find a line ***between*** two classes:

$$y = g(w_0 + w_1 x)$$



Linear Regression

Today's discussion: *Simple* Linear Regression

Given data $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$, $x_i, y_i \in \mathbb{R}$, we assume the data was sampled from a function:

$$y_i = f(x_i) + \epsilon_i$$

We wish to estimate this function as $\hat{f}(x_i)$ in order to predict,

$$\hat{y}_i = \hat{f}(x_i) \ .$$

In simple linear regression, we wish to estimate a line $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$. We can consider a choice of a *loss function* $(y_i - \hat{y}_i)^2$ that can inform us how well a certain estimate of $\hat{\beta}_0, \hat{\beta}_1$ is for the given data $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$.

Using calculus, we can arrive at the following closed-form expressions for $\hat{\beta}_0, \hat{\beta}_1$:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum_i x_i y_i - \bar{y} \sum_i x_i}{\sum_i x_i^2 - \bar{x} \sum_i x_i}$$

Programming Simple Linear Regression

1. Python Notebook: numpy-matplotlib.ipynb

2. Python Notebook: linear-regression-1.ipynb