

Human Roles in Machine Learning

Leon Tabak

Cornell College

07 October 2021

No more programming?

Pedro Domingos:

Machine learning is computers programming themselves instead of having to be programmed by us.

In the first stage of the information age we had to tell computers what to do, right?

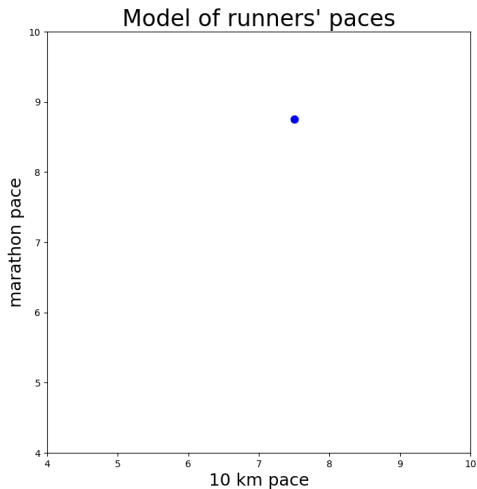
No more programming?

Pedro Domingos:

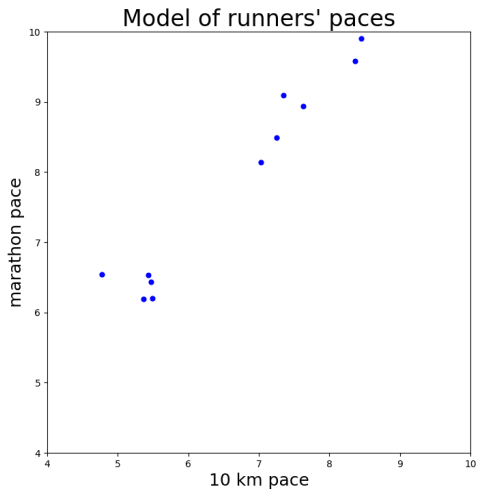
The new idea in machine learning is that we don't actually have to program the computers any more.

They look at the data that we generate and from that data, they try to figure out what to do.

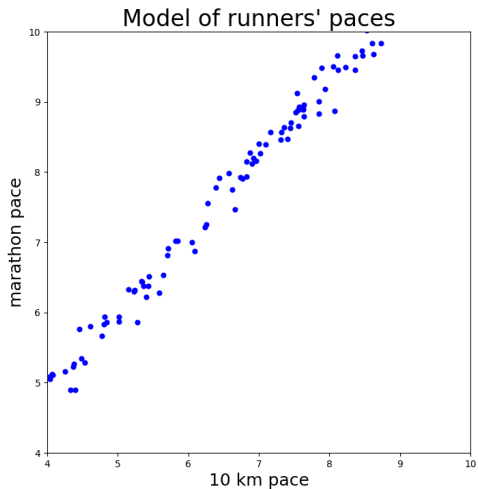
Method of Least Squares.



Method of Least Squares.



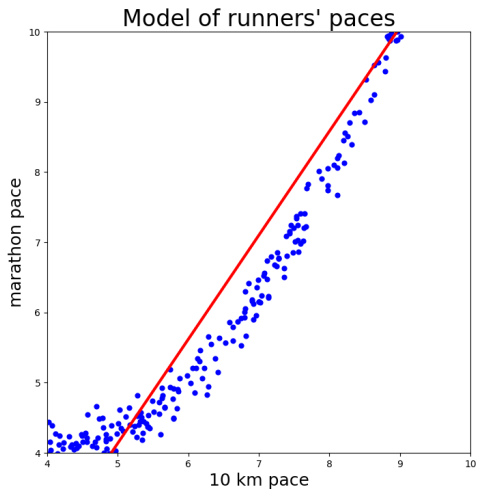
Method of Least Squares.



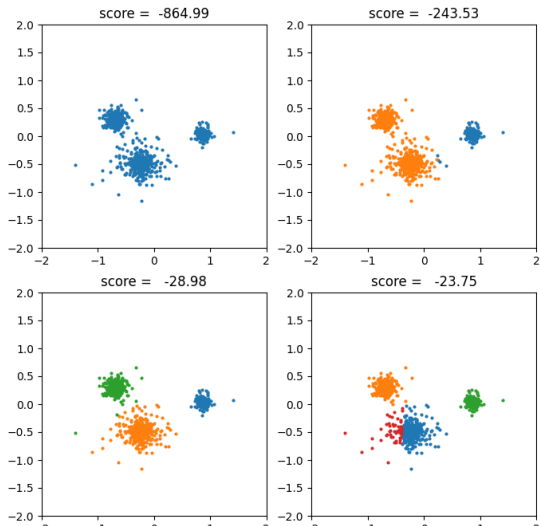
Method of Least Squares.



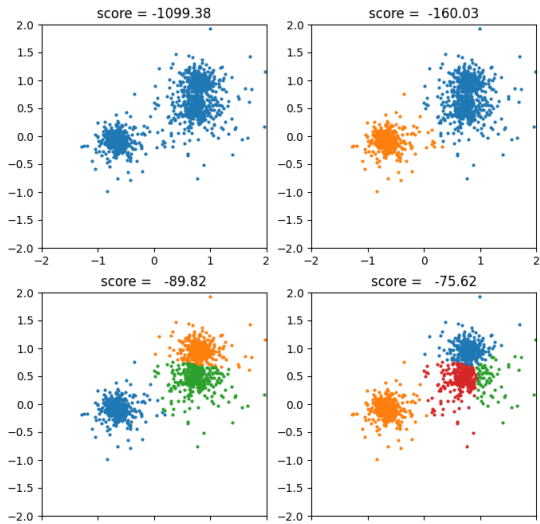
Method of Least Squares.



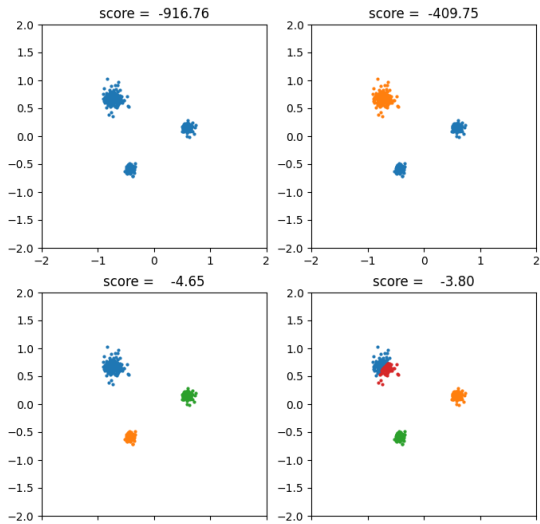
Clustering with K-Means



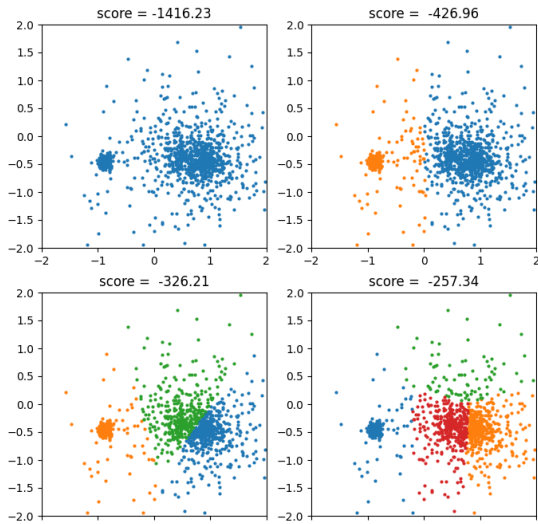
Clustering with K-Means



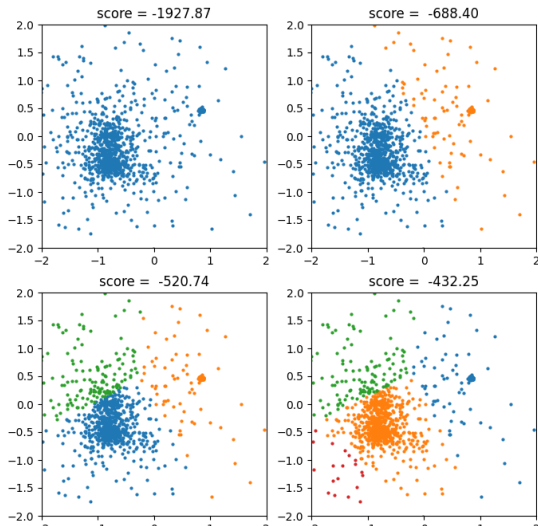
Clustering with K-Means



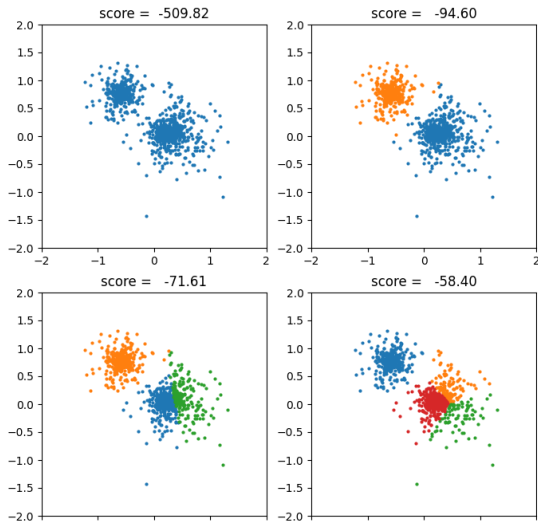
Clustering with K-Means



Clustering with K-Means



Clustering with K-Means



Machine learning today

Principal uses of machine learning. . .

- predication
- classification
- clustering

Machine learning today

Why has machine learning taken off in the last decade?

- it is easier to collect data
- it is easier to store data
- it is easier to share data
- it is easier to examine and analyze data

Machine learning today

Easier means. . .

- faster computations, faster data transmission
- smaller computing and storage devices
- less expensive computing, storage, and networking
- less special training needed

Machine learning today

- data-driven companies: Amazon, Facebook, Google, ...
- FOSS: free, open source software
- cloud computing

Machine learning's roots

Lagrange Interpolating Polynomial (1779, 1783, or 1795)

- Leonhard Euler (1707–1783)
- Joseph-Louis Lagrange (1736–1813)

Method of Least Squares (1794/95, 1805, or 1809)

- Carl Friedrich Gauss (1777–1855)
- Adrien-Marie Legendre (1752–1833)

Gradient Descent (1847, 1907)

- Augustin-Louis Cauchy (1789–1857)
- Jacques Hadamard (1865–1963)

Machine learning's roots

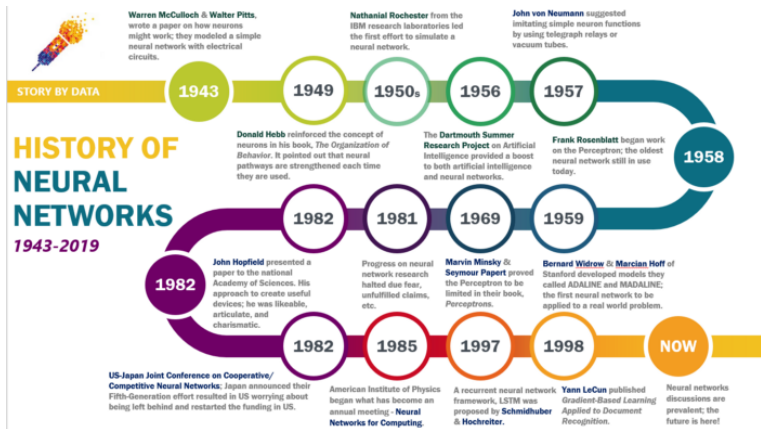
Principal Component Analysis (1901)

- Karl Pearson

K-Means (Lloyd's Algorithm) (1957)

- Stuart Lloyd, Bell Labs

Machine learning's roots



Machine learning's roots

We can solve a problem by guessing if we can. . .

- measure the error in each guess
- use that measure to make a better guess

“Learning representations by back-propagation”

David E. Rumelhart

Geoffrey E. Hinton

Ronald J. Williams

Nature, 323, 533–536 (1986)

How to improve a model

Any straight line in 2 dimensions can be described like this:

$$ax + by + c = 0$$

How to improve a model

We could improve our model by seeking a more complex relationship among the variables:

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

This model is no longer linear.

How to improve a model

Or we could stick with a linear model but add more variables...

$$aw + bx + cy + dz + e = 0$$