Perceptron

It all started with the perceptron...

Perceptron

Vol. VI, No. 2, Summer 1958

research trends

CORNELL AERONAUTICAL LABORATORY, INC., BUFFALO 21, NEW YORK

The Design of an Intelligent AUTOMATON

by FRANK ROSENBLATT

Introducing the perceptron — A machine which senses, recognizes, remembers, and responds like the human mind.

STORIES about the creation of machines having human qualities have long been a fascinating province in the realm of science fiction. Yet we are now about to witness the birth of such a machine — a machine capable of perceiving, recognizing, and identifying its surroundings without any human training or control.

Development of that machine has stemmed from a search for an understanding of the physical mechanisms which underlie human experience and intelligence. The question of the nature of these processes is at least as ancient as any other question in western science and philosophy, and, indeed, ranks as one of the greatest scientific challenges of our time.

Our understanding of this problem has gone perhaps as far as had the development of physics before Newton. We have some excellent descriptions of the phenomena to be explained, a number of interesting hypotheses, and a little detailed knowledge about events in the nervous system. But we lack agreement on any integrated set of principles by which the functioning of the nervous system can be understood.

We believe now that this ancient problem is about to yield to our theoretical investigation for three reasons:

First, in recent years our knowledge of the functioning of individual cells in the central nervous system has vastly increased.

Second, large numbers of engineers and mathematicians are, for the first time, undertaking serious study of the mathematical basis for thinking, perception, and the handling of information by the central nervous system, thus providing the hope that these problems may be within our intellectual grasp.

Third, recent developments in probability theory and in the mathematics of random processes provide new tools for the study of events in the nervous system, where only the gross statistical organization is known and the precise cell-by-cell "wiring diagram" may never be obtained.

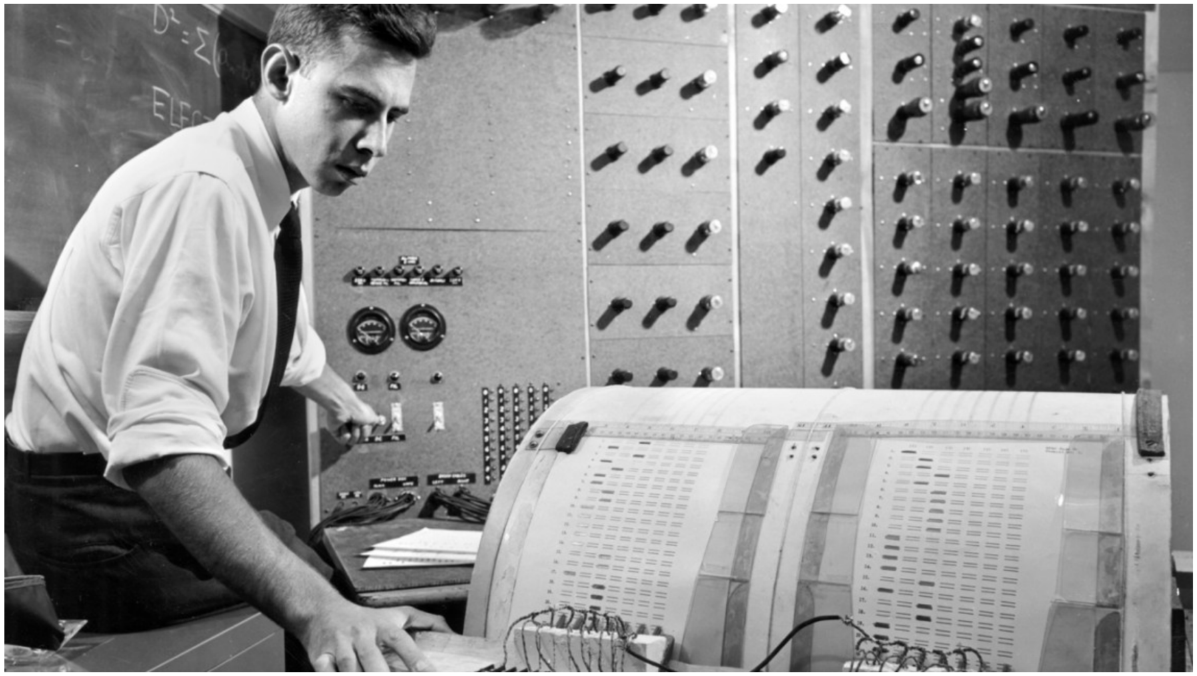
Receives Navy Support

In July, 1957, Project PARA (Perceiving and Recognizing Automaton), an internal research program which had been in progress for over a year at Cornell Aeronautical Laboratory, received the support of the Office of Naval Research. The program had been concerned primarily with the application of probability theory to

"An IBM 704 - a 5-ton computer the size of a room - was fed a series of punch cards. After 50 trials, the computer taught itself to distinguish cards marked on the left from cards marked on the right."

https://news.cornell.edu/stories/2019/09/professors-perceptron-paved-way-ai-60-years-too-soon

Perceptron



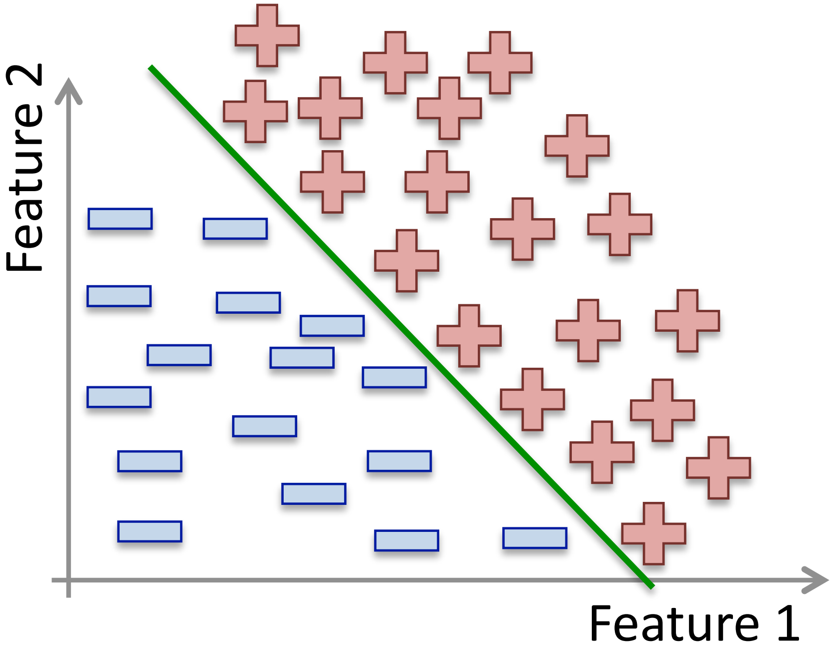
Division of Rare and Manuscript Collections

Frank Rosenblatt '50, Ph.D. '56, works on the "perceptron" – what he described as the first machine "capable of having an original idea."

Frank Rosenblatt's perceptron paved the way for AI 60 years ago.

Credit: https://news.cornell.edu/stories/2019/09/professors-perceptron-paved-way-ai-60-years-too-soon

Perceptron



Classification

Given:

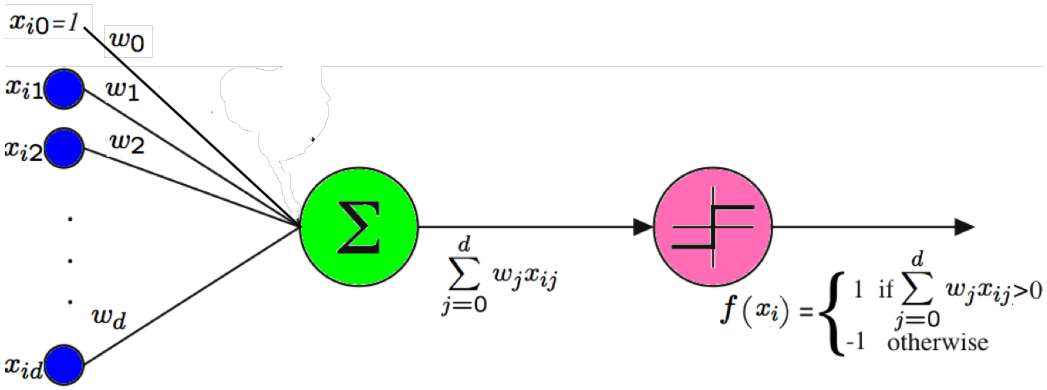
Training data: $(x_1, y_1), \ldots, (x_n, y_n) / x_i \in \mathbb{R}^d$ and y_i is discrete (categorical/qualitative), $y_i \in \mathbb{Y}$.

Task: Learn a classification function:

$$f : \mathbb{R}^d \longrightarrow \{-1, +1\}$$

$$f(x) = sign(\sum_{i=0}^{d} w_i x_i)$$

Perceptron


$x_{i0}=1$
w_0
x_{i1}
w_1
x_{i2}
w_2
w_d
x_{id}
Σ
$\sum_{j=0}^{d} w_j x_{ij}$
$f(x_i) = \begin{cases} 1 & \text{if } \sum_{j=0}^{d} w_j x_{ij} > 0 \\ -1 & \text{otherwise} \end{cases}$

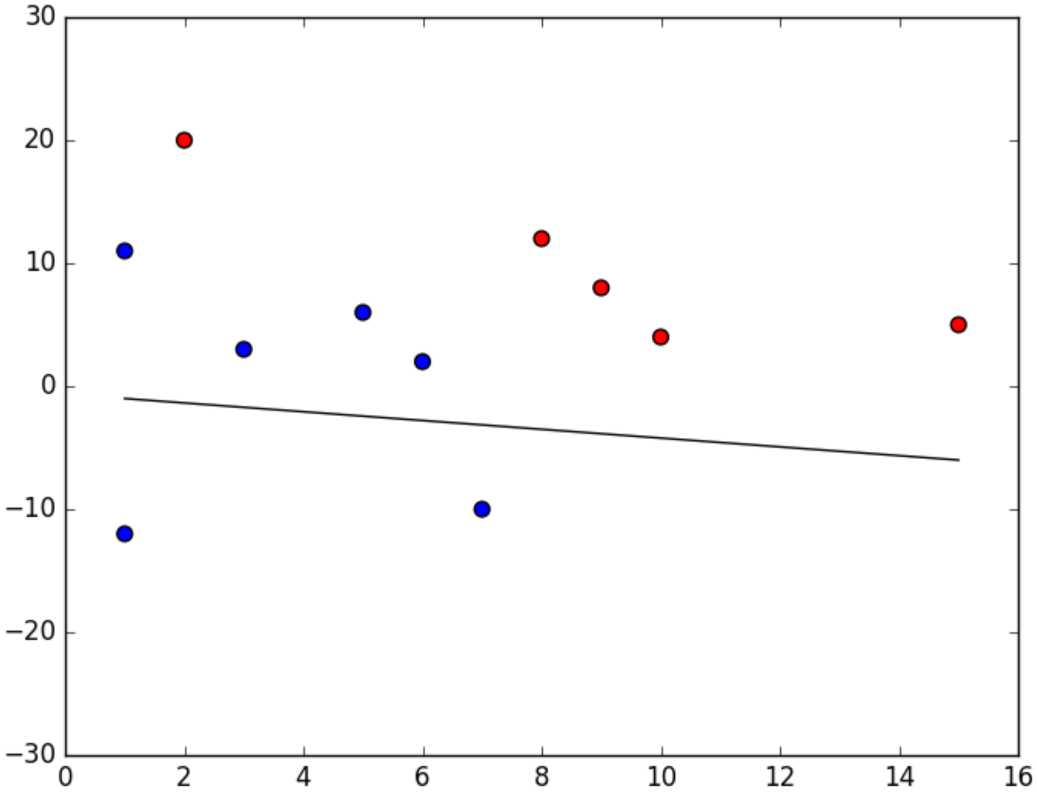
Given n examples and d features.

$$f(x_i) = sign(\sum_{j=0}^{d} w_j x_{ij})$$

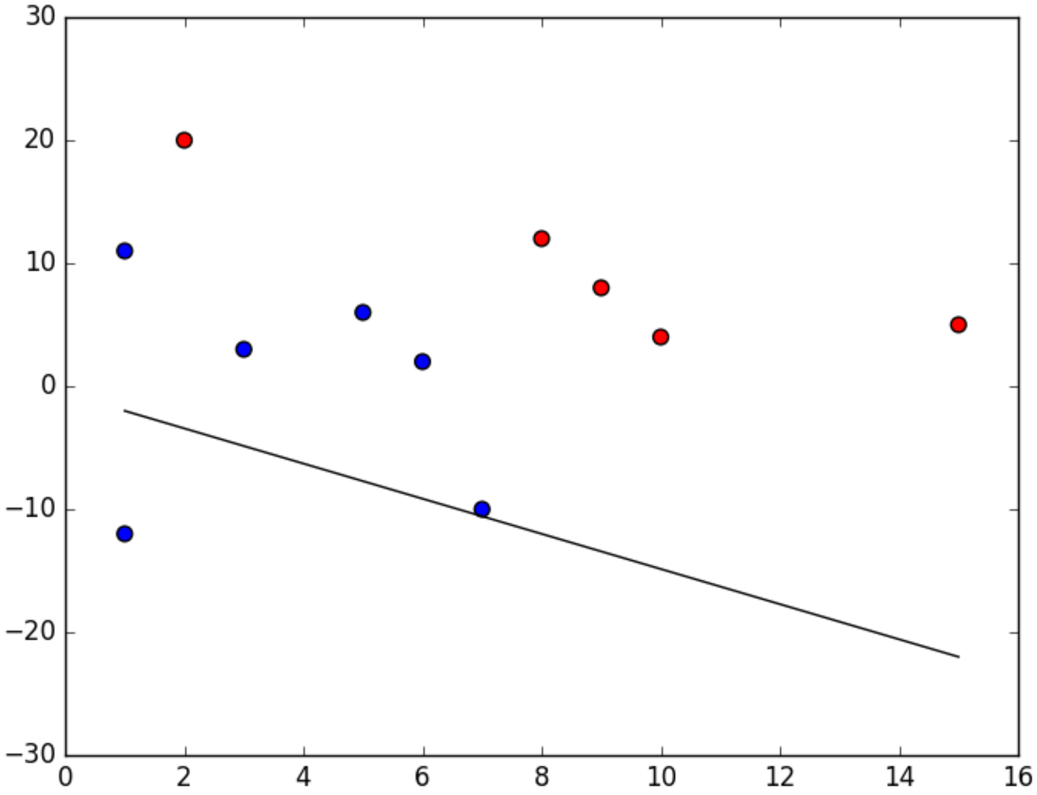
Perceptron

- Works perfectly if data is linearly separable. If not, it will not converge.
- Idea: Start with a random hyperplane and adjust it using your training data.
- Iterative method.

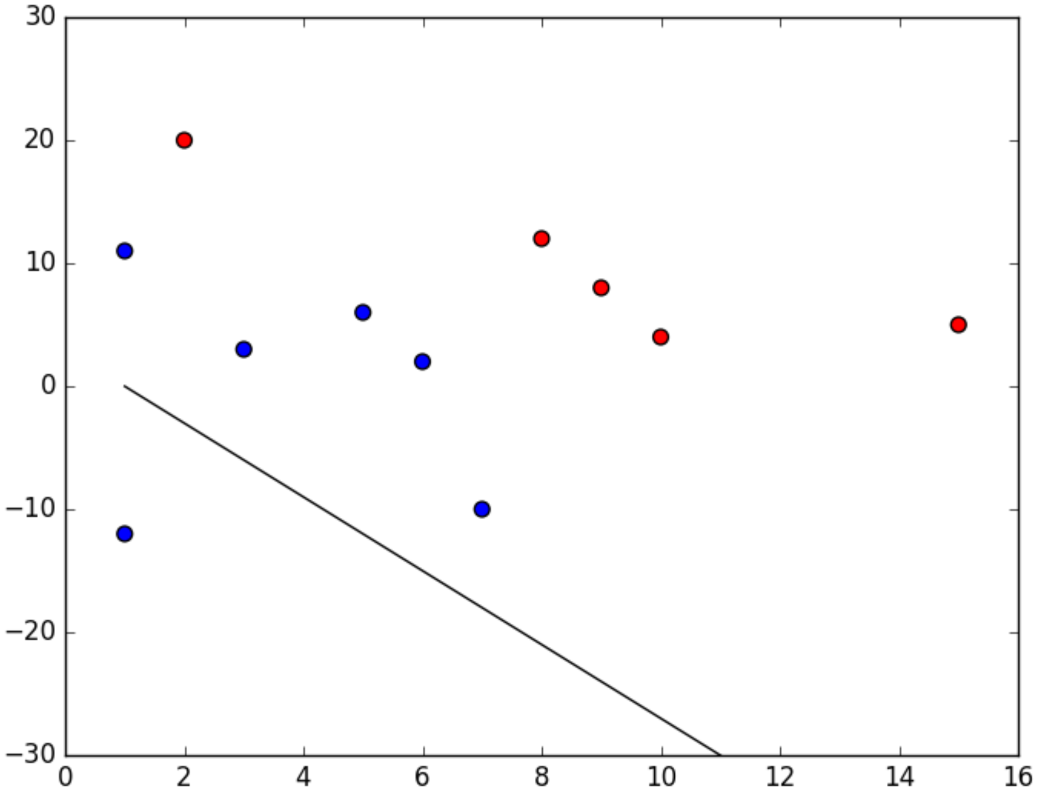
Perceptron: Example



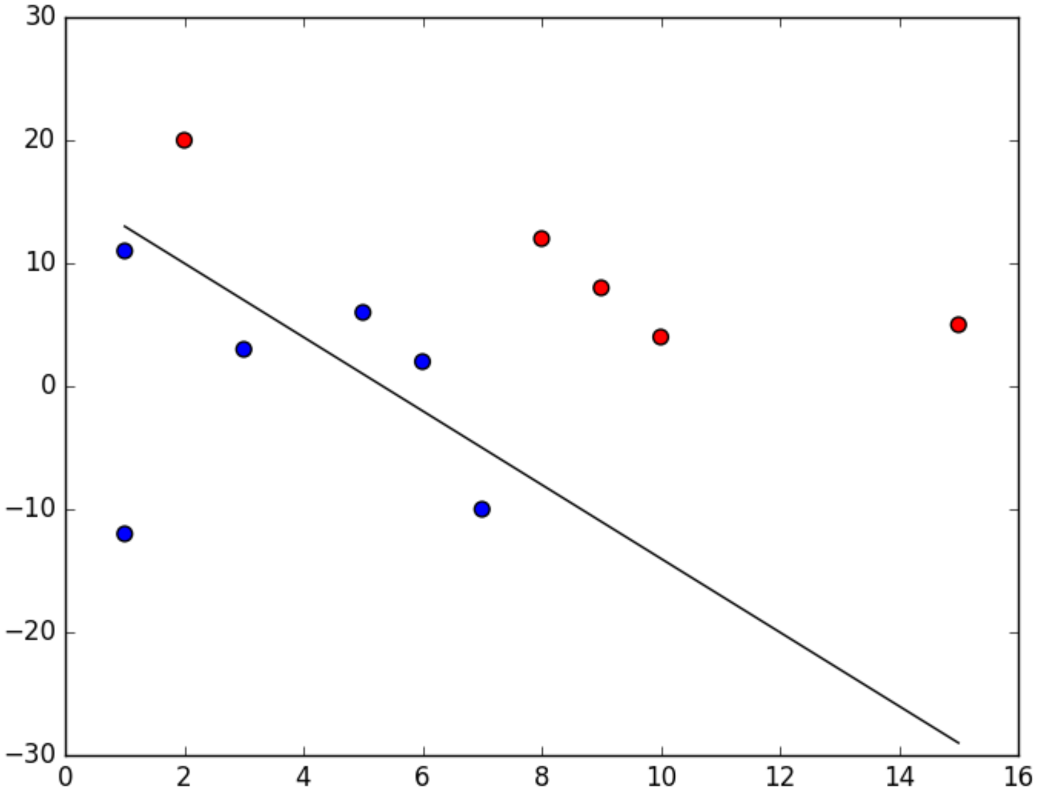
Perceptron: Example



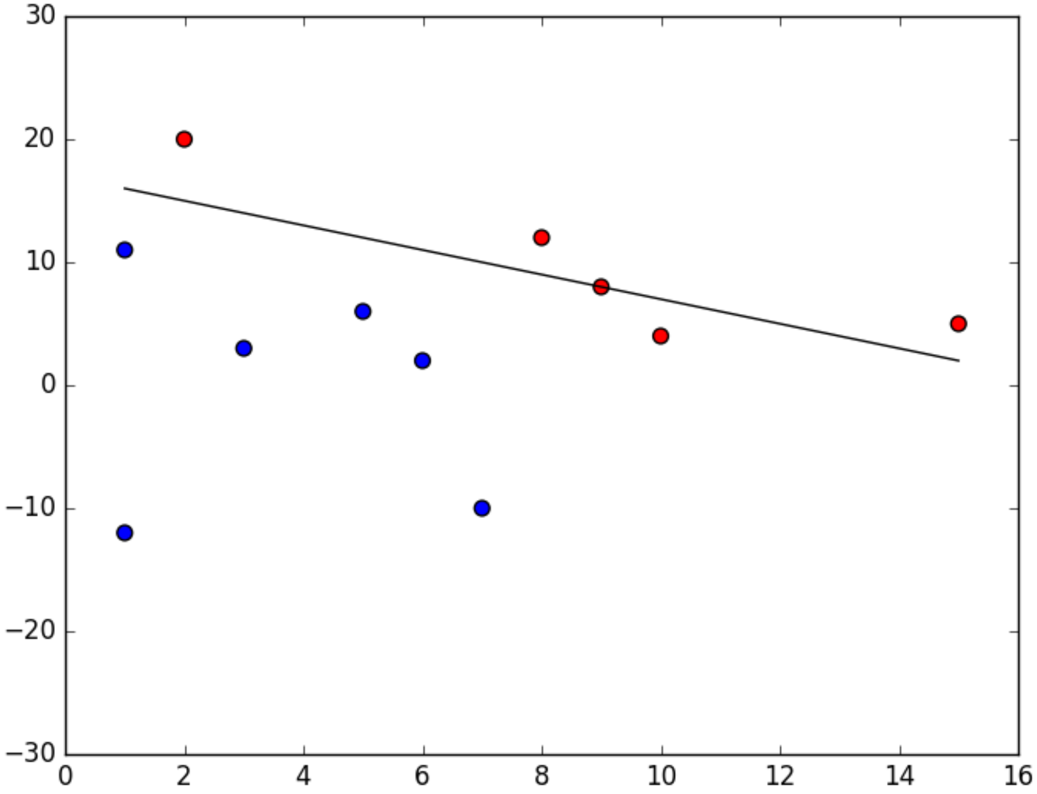
Perceptron: Example



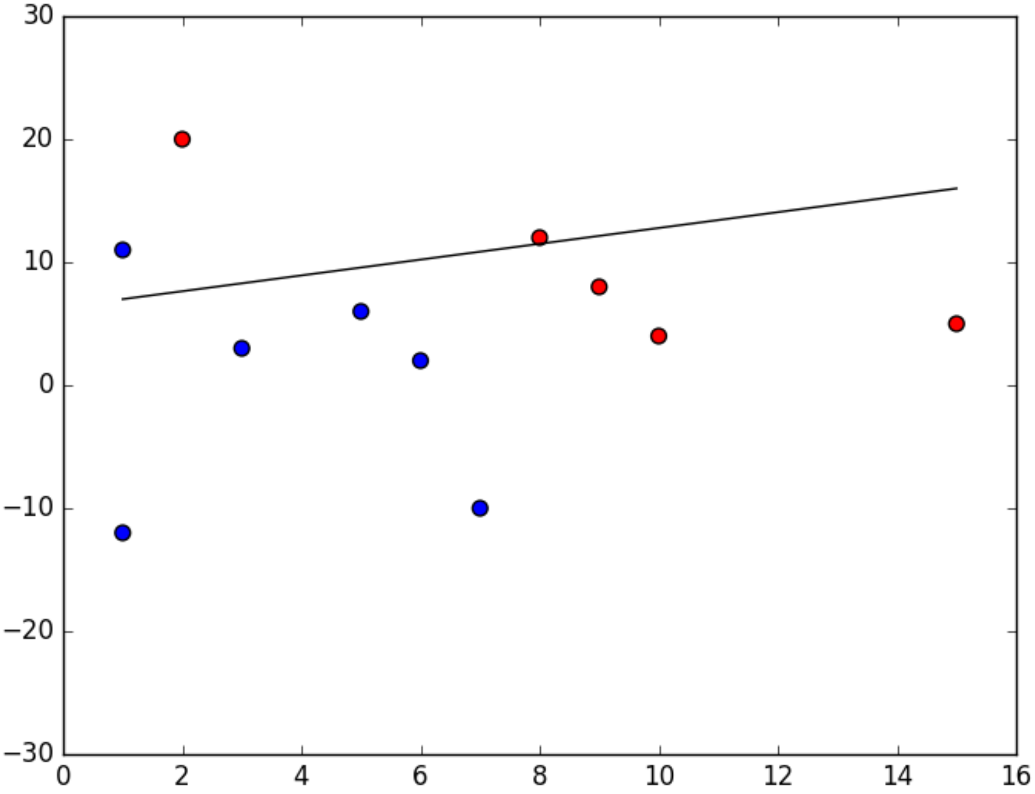
Perceptron: Example



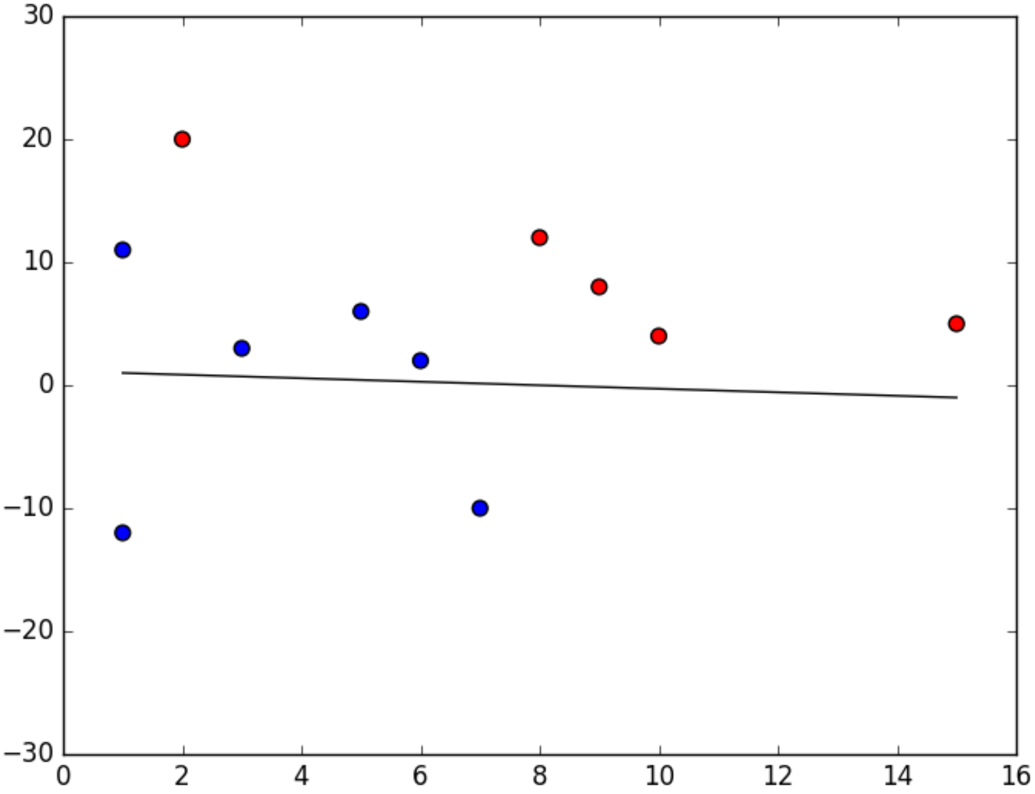
Perceptron: Example



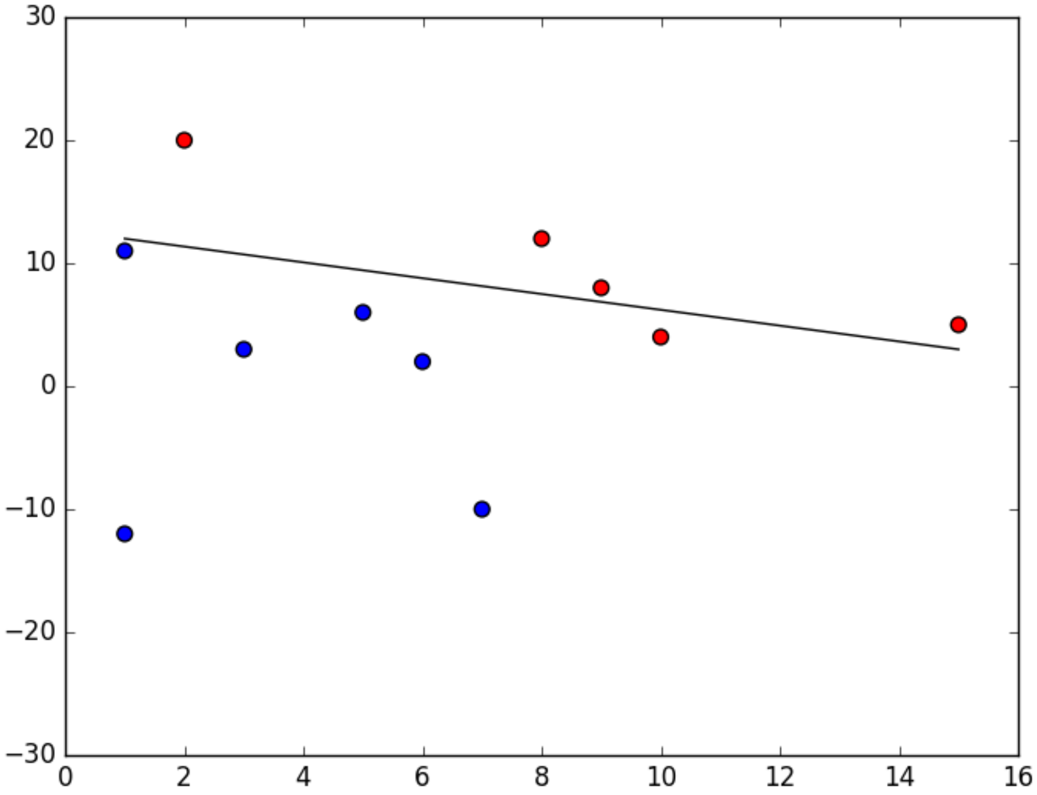
Perceptron: Example



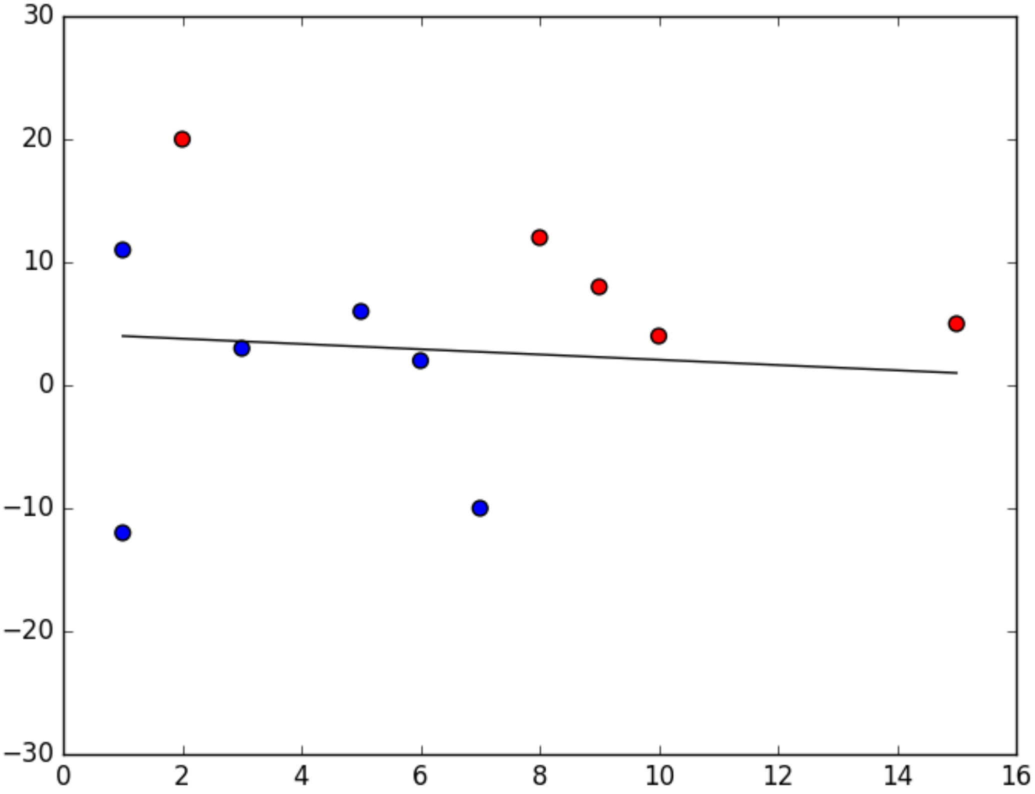
Perceptron: Example



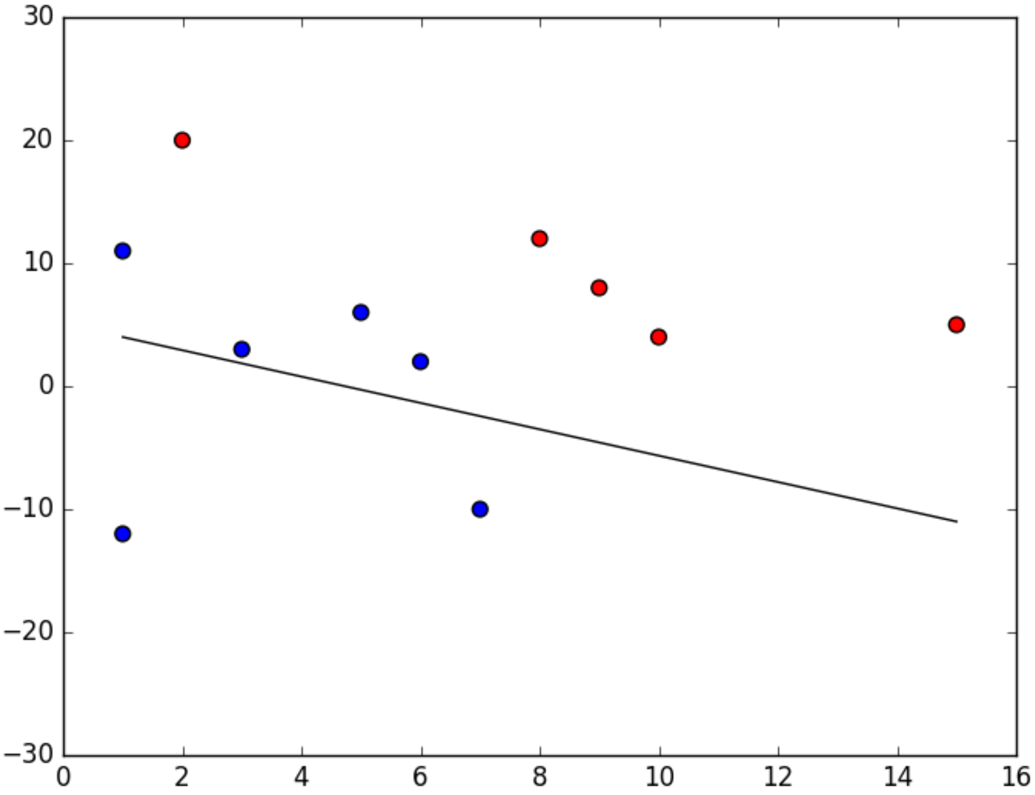
Perceptron: Example



Perceptron: Example

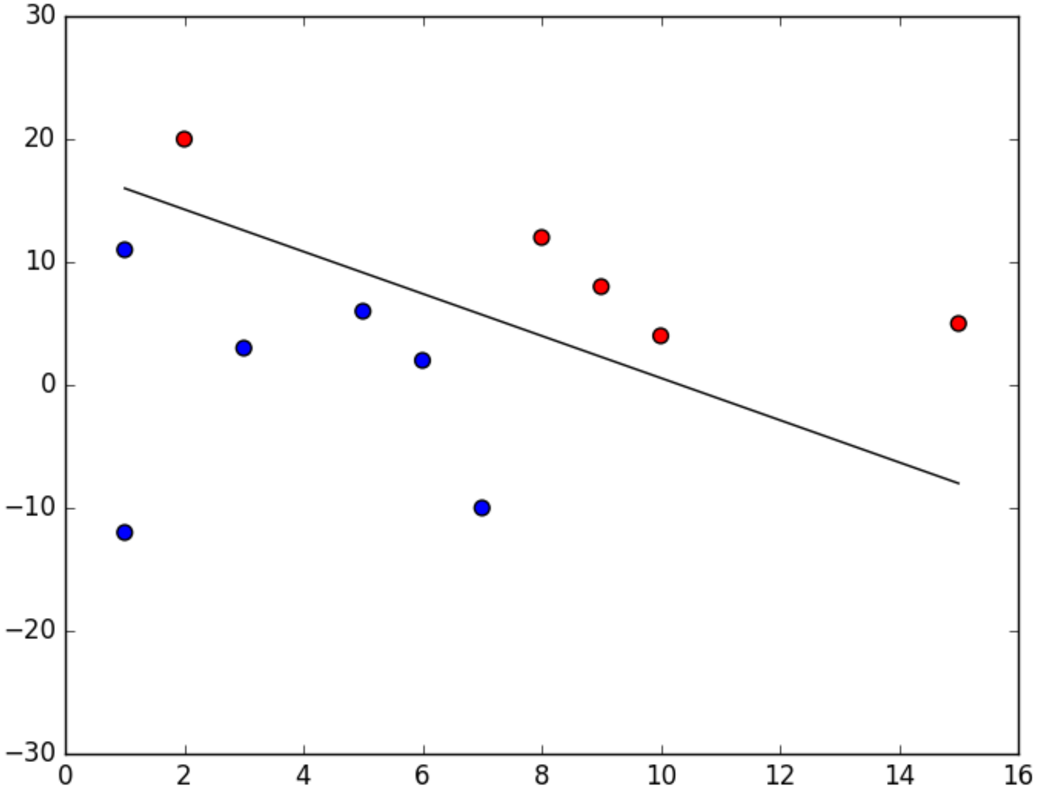


Perceptron: Example



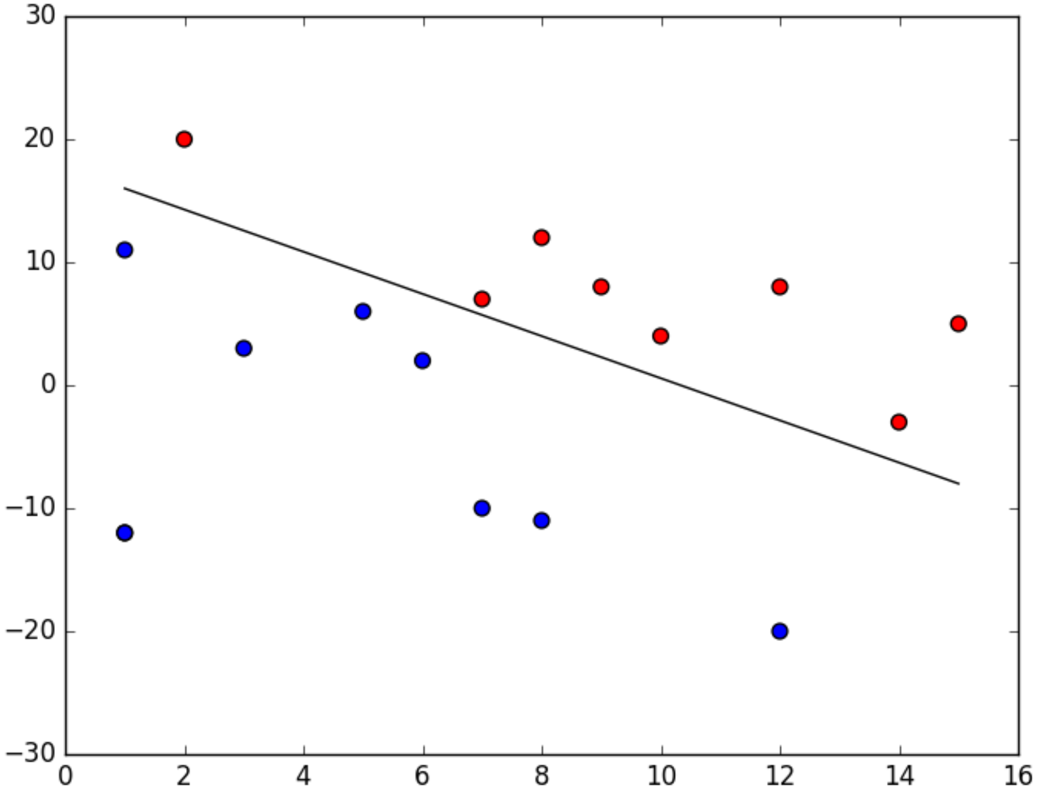
Perceptron: Example

Finally converged!

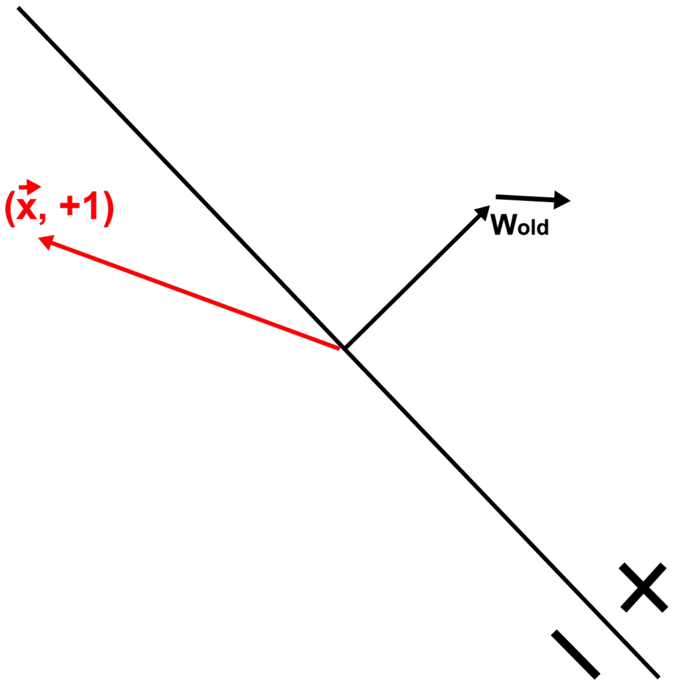


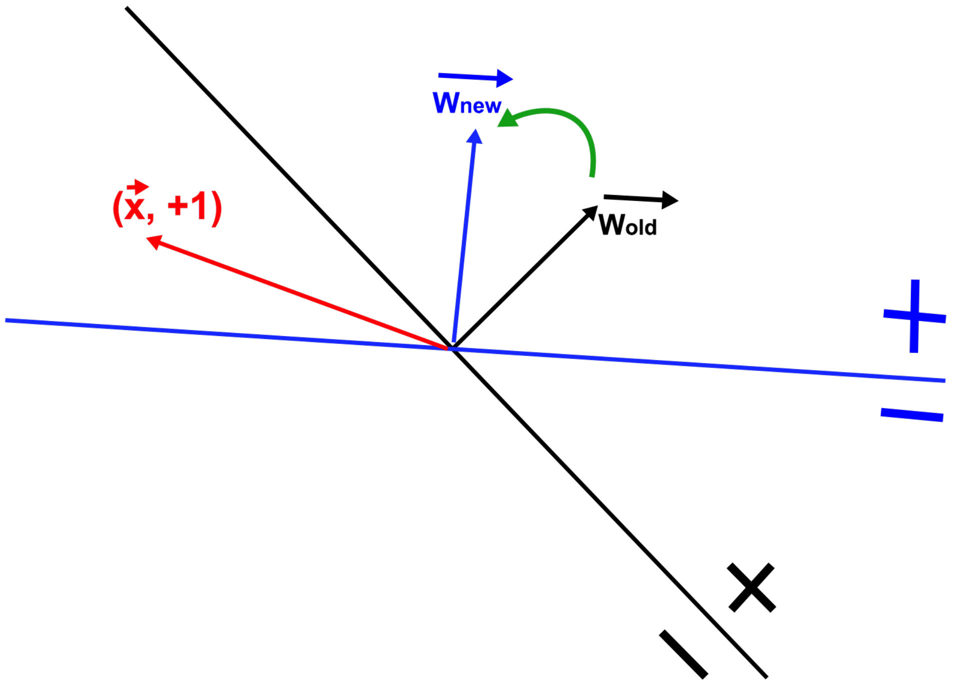
Perceptron: Example

With some test data:

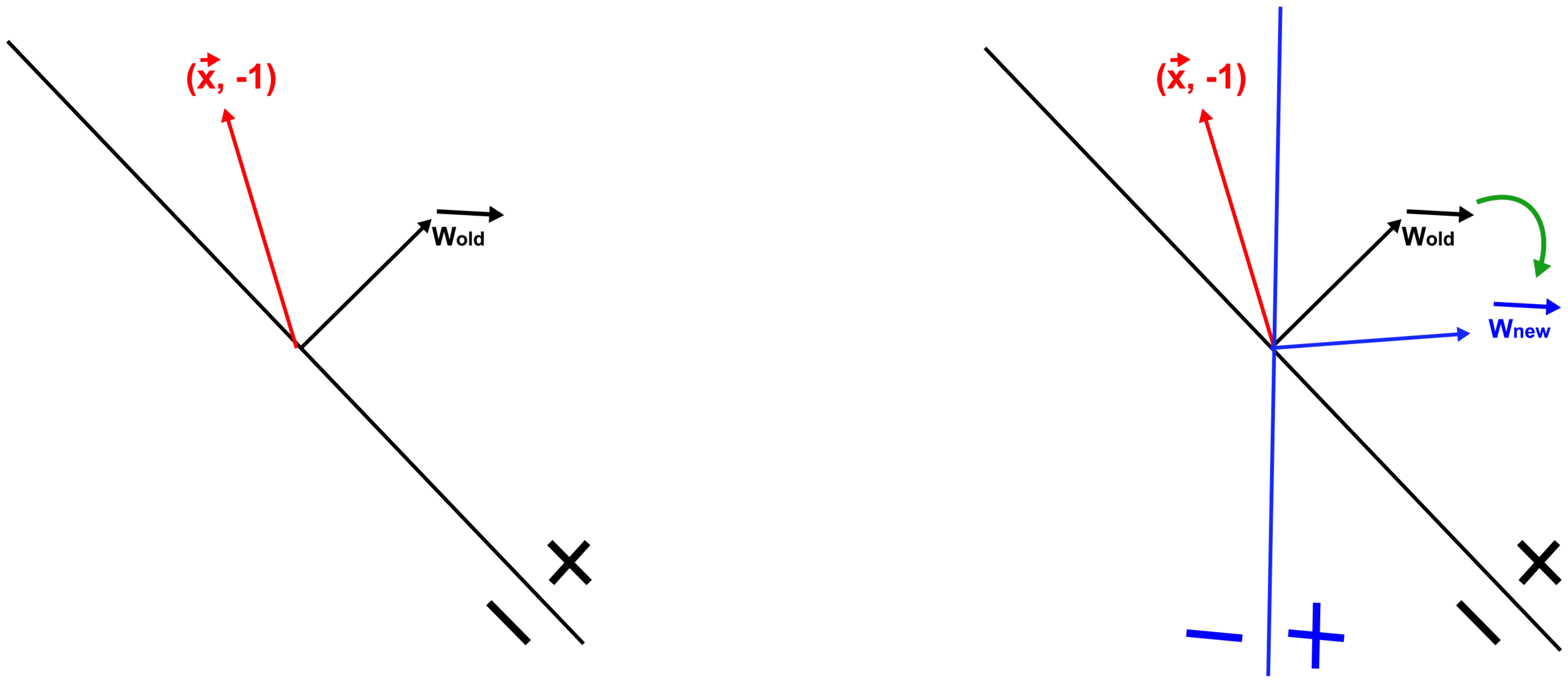


Perceptron visualization





Perceptron visualization



Perceptron expressiveness

- The perceptron works perfectly if data is linearly separable. If not, it will not converge.
- Idea: Iterative method that starts with a random hyperplane and adjust it using your training data.
- It can represent Boolean functions such as AND, OR, NOT but not the XOR function.
- Neural networks use the ability of the perceptrons to represent elementary functions and combine them in a network of layers of elementary questions: **Multiple Layer Perceptron (MLP)**.

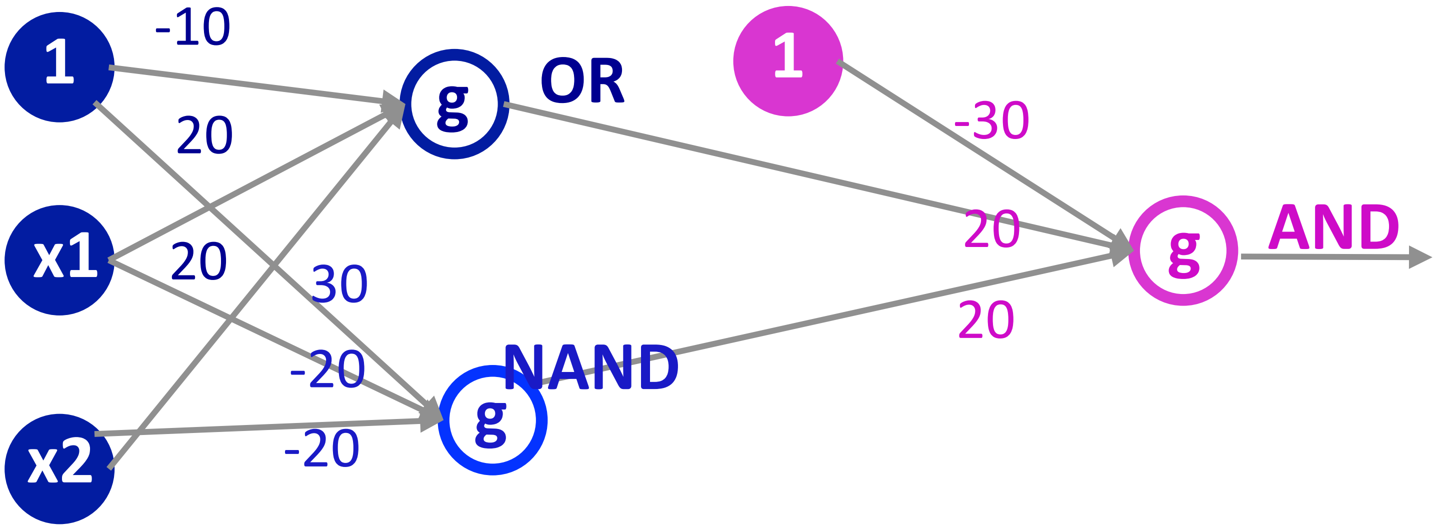
The XOR example

Let's try to create a MLP for the XOR function using elementary perceptrons.

x_1	x_2	x_1 **XOR** x_2	$(x_1$ **OR** $x_2)$ **AND** $(\ x_1$ **NAND** $x_2)$
0	0	0	0
0	1	1	1
1	0	1	1
1	1	0	0

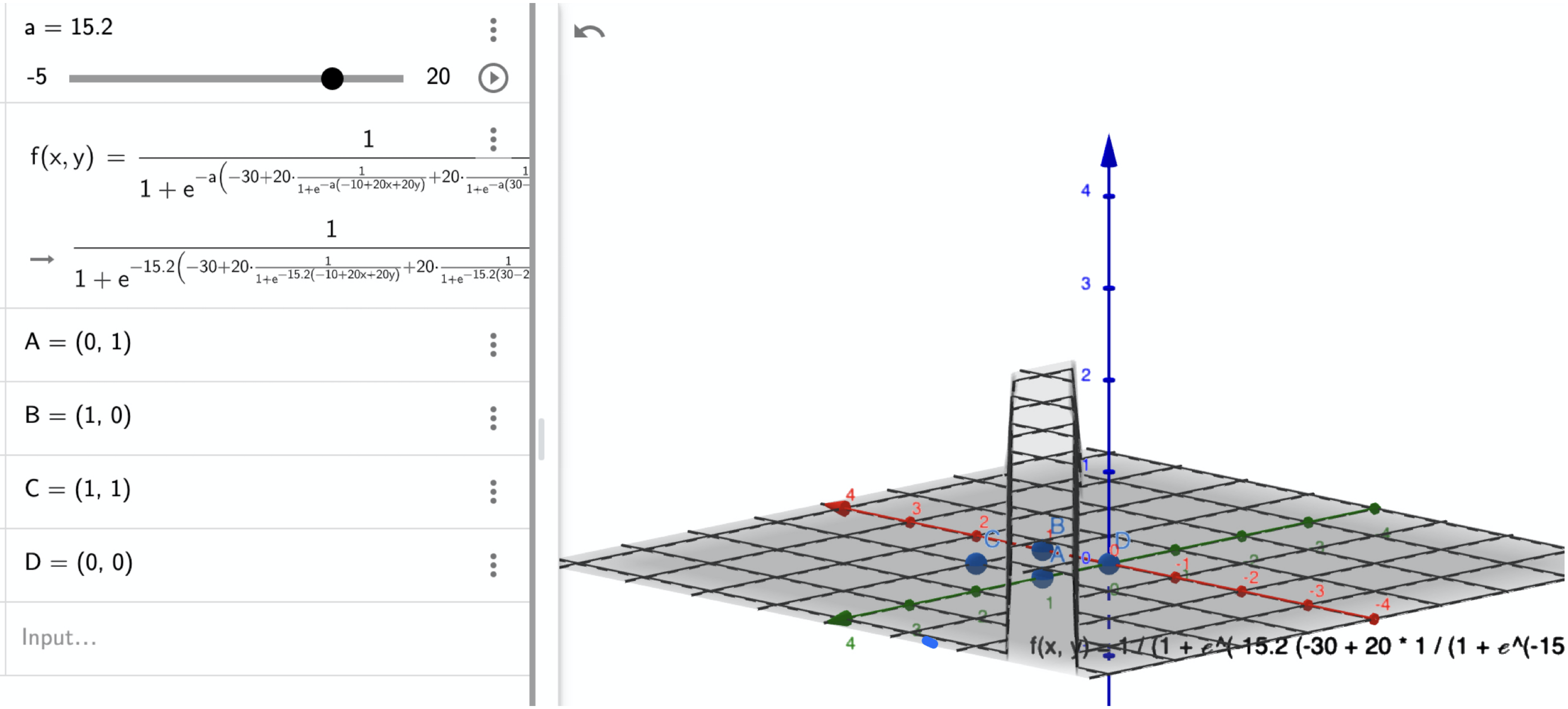
The XOR example

Let’s put them together...



XOR as a combination of 3 basic perceptrons.

The XOR example



a = 15.2

-5 ● 20

$$f(x, y) = \frac{1}{1 + e^{-a\left(-30+20\cdot\frac{1}{1+e^{-a(-10+20x+20y)}}+20\cdot\frac{1}{1+e^{-a(30-}}\right.}}$$

$$\rightarrow \frac{1}{1 + e^{-15.2\left(-30+20\cdot\frac{1}{1+e^{-15.2(-10+20x+20y)}}+20\cdot\frac{1}{1+e^{-15.2(30-2}}\right.}}$$

A = (0, 1)

B = (1, 0)

C = (1, 1)

D = (0, 0)

Input...

XOR as a combination of 3 basic perceptrons.