Introduction to Neural Networks and Deep Learning

P. Baldi

Department of Computer Science
Institute for Genomics and Bioinformatics
Center for Machine Learning and Intelligent Systems
University of California, Irvine
Intelligence in Brains and Machines

Deep Learning

Watson

Learning

Deep Learning
Early Beginnings

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

1. Completely Different Style of Computing
2. Universal Approximation Properties
Where do we store bits?
Synapses form at contact points between the dendrites of a cultured hippocampal neuron expressing YFP-actin (yellow) and axons from a CFP-actin expressing neurons (blue) whose cell bodies are present outside the field of view.
## Scales

<table>
<thead>
<tr>
<th></th>
<th>Size in Meters</th>
<th>$x10^6$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Diameter of Atom</td>
<td>$10^{-10}$</td>
<td>$10^{-4}$</td>
<td>Hair</td>
</tr>
<tr>
<td>Diameter of DNA</td>
<td>$10^{-9}$</td>
<td>$10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>Diameter of Synapse</td>
<td>$10^{-7}$</td>
<td>$10^{-1}$</td>
<td>Fist</td>
</tr>
<tr>
<td>Diameter of Axon</td>
<td>$10^{-6}$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Diameter of Neuron</td>
<td>$10^{-5}$</td>
<td>10</td>
<td>Room</td>
</tr>
<tr>
<td>Length of Axon</td>
<td>$10^{-3}$-$10^0$</td>
<td>$10^3$-$10^6$</td>
<td>Park-Nation</td>
</tr>
<tr>
<td>Length of Brain</td>
<td>$10^{-1}$</td>
<td>$10^5$</td>
<td>State</td>
</tr>
<tr>
<td>Length of Body</td>
<td>1</td>
<td>$10^6$</td>
<td>Nation</td>
</tr>
</tbody>
</table>
Let us assume that the persistence or repetition of a reverberatory activity (or “trace”) tends to induce lasting cellular changes that add to its stability…….*When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A’s efficiency, as one of the cells firing B, is increased.*

$$∆w_{ij} \sim O_i O_j$$
\[ a = W_1 x_1 + W_2 x_2 + \ldots + W_N x_N + \text{Bias} \]

\[ \text{output} = \begin{cases} 
-1, & \text{for all } a \leq 0 \\
1, & \text{for all } a > 0 
\end{cases} \]
Deep Learning Problem

$N_L$

$N_{L-1}$

$N_h$

$N_{h-1}$

$N_1$

$N_0$
BACK-PROPAGATION (~1985)

ERROR \( E(w) \)

\[
\Delta w_{ij} = \eta B_i O_j
\]
Many Engineering applications (computer vision, speech recognition, NLP, robotics, e-commerce, etc.)
Similar trends across engineering disciplines from speech recognition, to NLP, to games, and to robotics.

“super-human” performance in 2015
Natural Sciences: Yesterday

Galileo Galilei (1564-1642)
Natural Sciences: Today

5000x
Application: Breast Arterial Calcification Detection

Fig. 1 (left) A sample mammogram ROI containing BACs, (right) detection results by the deep neural network, in which detections are indicated by the green color and human marked BACs are indicated by red color.

Fig. 2. FROC curves obtained by the deep neural network.

Fig. 3. Calcium mass calculated from the human marked BACs (Reference Ca mass) and from detections (Calculated Ca mass).
THANK YOU