Lecture 23: Perceptrons and their computing power

8th March, 2011

(Lectures 21 and 22 were on Text Entailment by Prasad Joshi)
A perspective of AI
Artificial Intelligence - Knowledge based computing
Disciplines which form the core of AI - inner circle
Fields which draw from these disciplines - outer circle.
Neuron - “classical”

- **Dendrites**
  - Receiving stations of neurons
  - Don't generate action potentials

- **Cell body**
  - Site at which information received is integrated

- **Axon**
  - Generate and relay action potential
  - Terminal
    - Relays information to next neuron in the pathway

Computation in Biological Neuron

- Incoming signals from synapses are summed up at the soma $\sum$, the biological “inner product”
- On crossing a threshold, the cell “fires” generating an action potential in the axon hillock region

Synaptic inputs: Artist’s conception
The Perceptron Model

A perceptron is a computing element with input lines having associated weights and the cell having a threshold value. The perceptron model is motivated by the biological neuron.
Step function / Threshold function

\[ y = 1 \text{ for } \sum w_i x_i \geq \theta \]
\[ = 0 \text{ otherwise} \]
Features of Perceptron

• Input output behavior is discontinuous and the derivative does not exist at $\sum w_i x_i = \theta$

• $\sum w_i x_i - \theta$ is the net input denoted as net

• Referred to as a linear threshold element - linearity because of $x$ appearing with power $1$

• $y = f(\text{net})$: Relation between $y$ and net is non-linear
Computation of Boolean functions

AND of 2 inputs

<table>
<thead>
<tr>
<th>X1</th>
<th>x2</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

The parameter values (weights & thresholds) need to be found.
Computing parameter values

\[ w_1 \times 0 + w_2 \times 0 \leq \theta \Rightarrow \theta \geq 0; \text{ since } y=0 \]

\[ w_1 \times 0 + w_2 \times 1 \leq \theta \Rightarrow w_2 \leq \theta; \text{ since } y=0 \]

\[ w_1 \times 1 + w_2 \times 0 \leq \theta \Rightarrow w_1 \leq \theta; \text{ since } y=0 \]

\[ w_1 \times 1 + w_2 \times 1 > \theta \Rightarrow w_1 + w_2 > \theta; \text{ since } y=1 \]

\[ w_1 = w_2 = 0.5 \]

satisfy these inequalities and find parameters to be used for computing AND function.
Other Boolean functions

• OR can be computed using values of \( w_1 = w_2 = 1 \) and \( \theta = 0.5 \)

• XOR function gives rise to the following inequalities:
  \[
  \begin{align*}
    w_1 \cdot 0 + w_2 \cdot 0 & \leq \theta \Rightarrow \theta \geq 0 \\
    w_1 \cdot 0 + w_2 \cdot 1 & > \theta \Rightarrow w_2 > \theta \\
    w_1 \cdot 1 + w_2 \cdot 0 & > \theta \Rightarrow w_1 > \theta \\
    w_1 \cdot 1 + w_2 \cdot 1 & \leq \theta \Rightarrow w_1 + w_2 \leq \theta 
  \end{align*}
  \]

No set of parameter values satisfy these inequalities.
Threshold functions

<table>
<thead>
<tr>
<th>n</th>
<th># Boolean functions ((2^{2^n}))</th>
<th>#Threshold Functions ((2^{n^2}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>256</td>
<td>128</td>
</tr>
<tr>
<td>4</td>
<td>64K</td>
<td>1008</td>
</tr>
</tbody>
</table>

- Functions computable by perceptrons - threshold functions
- #TF becomes negligibly small for larger values of #BF.
- For \(n=2\), all functions except XOR and XNOR are computable.
Concept of Hyper-planes

- $\sum w_i x_i = \theta$ defines a linear surface in the $(W, \theta)$ space, where $W=\langle w_1, w_2, w_3, \ldots, w_n \rangle$ is an $n$-dimensional vector.

- A point in this $(W, \theta)$ space defines a perceptron.
**Perceptron Property**

- Two perceptrons may have different parameters but same functional values.

- Example of the simplest perceptron
  - $w \cdot x > 0$ gives $y = 1$
  - $w \cdot x \leq 0$ gives $y = 0$

Depending on different values of $w$ and $\theta$, four different functions are possible.
Simple perceptron contd.

<table>
<thead>
<tr>
<th>x</th>
<th>f1</th>
<th>f2</th>
<th>f3</th>
<th>f4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

0-function

- $\theta \geq 0$
- $w \leq 0$

Identity Function

- $\theta \geq 0$
- $w > 0$

Complement Function

- $\theta < 0$
- $w \leq 0$

True-Function

- $\theta < 0$
- $W < 0$
Counting the number of functions for the simplest perceptron

- For the simplest perceptron, the equation is \( w \cdot x = \theta \).

Substituting \( x = 0 \) and \( x = 1 \), we get \( \theta = 0 \) and \( w = \theta \).

These two lines intersect to form four regions, which correspond to the four functions.
Fundamental Observation

- The number of TFs computable by a perceptron is equal to the number of regions produced by $2^n$ hyper-planes, obtained by plugging in the values $<x_1,x_2,x_3,...,x_n>$ in the equation

$$\sum_{i=1}^{n} w_i x_i = \theta$$
The geometrical observation

**Problem:** m linear surfaces called hyperplanes (each hyper-plane is of (d-1)-dim) in d-dim, then what is the max. no. of regions produced by their intersection? i.e. $R_{m,d} = ?$
Co-ordinate Spaces

We work in the \( <X_1, X_2> \) space or the \( <w_1, w_2, \theta> \) space

General equation of a Hyperplane:
\[
\Sigma W_i X_i = \theta
\]

Hyperplane (Line in 2-D):
\[
W_1 = W_2 = 1, \quad \theta = 0.5 \\
X_1 + x_2 = 0.5
\]
Regions produced by lines

Regions produced by lines not necessarily passing through origin
L1: 2
L2: 2+2 = 4
L2: 2+2+3 = 7
L2: 2+2+3+4 =
11

New regions created = Number of intersections on the incoming line by the original lines
Total number of regions = Original number of regions + New regions created
Number of computable functions by a neuron

\[ w_1 * x_1 + w_2 * x_2 = \theta \]

\((0,0) \Rightarrow \theta = 0 : P1\)

\((0,1) \Rightarrow w_2 = \theta : P2\)

\((1,0) \Rightarrow w_1 = \theta : P3\)

\((1,1) \Rightarrow w_1 + w_2 = \theta : P4\)

P1, P2, P3 and P4 are planes in the \(<W_1,W_2, \theta>\) space
Number of computable functions by a neuron (cont...)

- P1 produces 2 regions
- P2 is intersected by P1 in a line. 2 more new regions are produced.
  Number of regions = 2 + 2 = 4
- P3 is intersected by P1 and P2 in 2 intersecting lines. 4 more regions are produced.
  Number of regions = 4 + 4 = 8
- P4 is intersected by P1, P2 and P3 in 3 intersecting lines. 6 more regions are produced.
  Number of regions = 8 + 6 = 14
- Thus, a single neuron can compute 14 Boolean functions which are linearly separable.
Points in the same region

If
\[ W_1 X_1 + W_2 X_2 > \theta \]
\[ W_1' X_1 + W_2' X_2 > \theta' \]
Then
If \(<W_1, W_2, \theta>\) and \(<W_1', W_2', \theta'>\) share a region then they compute the same function.