LA-UR- 97-60

CONF-9706117--1

Title: TIGHT BOUNDS ON THE SIZE OF NEURAL NETWORKS
FOR CLASSIFICATION PROBLEMS

Author(s): Valeriu Beiu
Thierry de Pauw

Submitted to: International Work-Conference on Artificial and Natural
Neural Networks
June 4-6, 1997
Lanzarote-Canary Islands, Spain

Los Alamos
NATIONAL LABORATORY

DISCLAIMER

Tight Bounds on the Size of Neural Networks for Classification Problems

Valeriu Beiu [1, *]   **and   Thierry de Pauw** [2]

[1] Los Alamos National Laboratory
Division NIS-1
Mail Stop D466
Los Alamos, NM 87545, USA

[2] Université Catholique de Louvain
Départment de Mathématique
Chemin du Cyclotron, 2
B-1348 Louvain-la-Neuve, Belgium

Abstract—This paper relies on the entropy of a data-set (i.e., *number-of-bits*) to prove tight bounds on the *size* of neural networks solving a classification problem. First, based on a sequence of geometrical steps, we constructively compute an upper bound of $O(mn)$ on the *number-of-bits* for a given data-set—here m is the number of examples and n is the number of dimensions (i.e., $\mathbb{R}^n$). This result is used further in a nonconstructive way to bound the *size* of neural networks which correctly classify that data-set.

Keywords—neural networks, *size* complexity, entropy, classification problems, Boolean functions.

1. Introduction

Multilayer feedforward neural networks (NNs) have been experimentally shown to be quite effective in many different applications (see *Applications of Neural Networks* in [3], together with *Part F: Applications of Neural Computation* and *Part G: Neural Networks in Practice: Case Studies* from [13]). This success has led researchers to undertake a rigorous analysis of the mathematical properties that enable them to perform so well, and has generated two directions of research:

- one to find existence /constructive proofs for what is now known as the *"universal approximation problem"* (i.e., any continuous function can be approximated arbitrarily well by a NN);

- another one to find tight bounds on the number of neurons (*size*) needed by the approximation problem (or some particular cases).

This paper will prove several bounds—based on the entropy of the data-set—on the *size* of NNs for classification problems. In Section 2 we introduce several notations and in Section 3 we briefly go through some pervious results. The proof will be given in Section 4. It is based on computing the *number-of-bits* needed for representing the data-set as the quotient between the volume of a ball containing all the examples (rough approximation) and the volume of a minimum hypercube [6]. Because all the examples from one class always lie inside the intersection of two balls, we shall compute the volume of this intersection. Using this volume it can be proven that the required *number-of-bits* is slightly lower. Finally, based on previous bounds on the magnitude of the *weights* of threshold gates, we shall bound the *size* of NNs solving the classification problem specified by the data-set.

* On leave of absence from the "Politehnica" University of Bucharest, Department of Computer Science, Spl. Independentei 313, RO–77206 Bucharest, Romania.

2. Several Notations

In this paper a *network* is an acyclic graph having several input nodes (*inputs*) and some (at least one) output nodes (*outputs*). If with each edge a synaptic *weight* is associated and each node computes the weighted sum of its inputs to which a nonlinear activation function is then applied (*artificial neuron*):

$$f(z) = f(z_1, ..., z_\Delta) = \sigma\left(\sum_{i=1}^{\Delta} w_i z_i + \theta\right), \tag{1}$$

the network is a NN ($w_i \in \mathbb{R}$ are the synaptic *weights*, $\theta \in \mathbb{R}$ is the *threshold*, Δ is the number of inputs of one neuron, and σ is a non-linear activation function).

A *classification problem* is defined by a set of m examples (i.e., data-set) belonging to k different classes. For simplicity we shall limit the number of classes to two ($k = 2$), known as a *dichotomy*, but all our results are valid in general. Now:

$$m = m_+ + m_- \tag{2}$$

and $x_1, x_2, ..., x_{m_+}$ are the positive examples, while $y_1, y_2, ..., y_{m_-}$ are the negative examples; they are taken from an n-dimensional space $\mathbb{R}^n$ ($n \in \mathbb{N} \setminus \{1\}$):

$$x_i = (x_{i,1}, x_{i,2}, ..., x_{i,n}) \in \mathbb{R}^n, i = 1, 2, ..., m_+, m_+ \in \mathbb{N} \text{ and}$$

$$y_j = (y_{j,1}, y_{j,2}, ..., y_{j,n}) \in \mathbb{R}^n, j = 1, 2, ..., m_-, m_- \in \mathbb{N}. \tag{3}$$

The *distance* between two vectors (examples) is the classical *Euclidean distance*:

$$dist_E(a, b) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + ... + (a_n - b_n)^2} = \left\{\sum_{i=1}^{n}(a_i - b_i)^2\right\}^{1/2}. \tag{4}$$

For characterising the data-set, we also define the *minimum* and the *maximum distance* between any positive and negative examples:

$$d = \min_{\substack{i=1,...,m_+ \\ j=1,...,m_-}} [dist_E(x_i, y_j)] \text{ and } D = \max_{\substack{i=1,...,m_+ \\ j=1,...,m_-}} [dist_E(x_i, y_j)] \tag{5}$$

A *ball* of radius r ($r \in \mathbb{R}^+ \setminus \{0\}$) centered at $c \in \mathbb{R}^n$ will be denoted by $B_n[c, r]$:

$$B_n[c, r] = \{x \in \mathbb{R}^n : dist_E(c, x) \le r\}; \tag{6}$$

if $n = 2$ this is a round disc; if $n = 3$ it is a round ball.

We shall denote by μ_n the *n-dimensional Lebesgue measure* in $\mathbb{R}^n$. If $A \subseteq \mathbb{R}^2$, $\mu_2(A)$ is the '*area*' of A; if $A \subseteq \mathbb{R}^3$, $\mu_3(A)$ is the '*volume*' of A.

Finally, $\alpha(n) = \mu_n(B_n[0, 1])$ is the *volume of the unit ball* in $\mathbb{R}^n$. For instance $\alpha(2) = \pi$, $\alpha(3) = 4\pi/3$, while in general $\alpha(2n-1) = 2^n \cdot \pi^{n-1}/[1\cdot3\cdot5\cdot...\cdot(2n-1)]$ and $\alpha(2n) = \pi^n/n!$ [11, 15, 17], or in terms of the gamma function:

$$\alpha(n) = \frac{\pi^{n/2}}{\Gamma(n/2 + 1)}. \tag{7}$$

3. Previous Results

The problem to find the smallest *size* NN which can realize an arbitrary function given a set of m vectors (examples, or points) in n dimensions is not new. Many results have been obtained for NNs having a *threshold* activation function. This is probably due to the fact that this line of research was continuing on the rigorous results already obtained in the literature dealing with threshold logic from the mid 60s (see references in [7, 8]). Probably the first lower bound on the *size* of a threshold gate circuit for "almost all" n-ary Boolean functions was given by Neciporuk in 1964: $size \geq 2 \, (2^n/n)^{1/2}$ [20]. Later, Lupanov has proven a very tight upper bound: $size \leq 2 \, (2^n/n)^{1/2} \times \{1 + \Omega \, [(2^n/n)^{1/2}]\}$ for the case when $depth = 4$ [19]. Similar existence exponential bounds can be found in [12], while in [24] a $\Omega \, (2^{n/3})$ existence lower bound for arbitrary Boolean functions has been presented.

For classification problems, one of the first result was that a NN with only one hidden layer having $m - 1$ nodes could compute an arbitrary dichotomy (sufficient condition). The main improvements have been as follows:

- Baum [4] presented a NN with one hidden layer having $\lceil m/n \rceil$ neurons [1] capable of realizing an arbitrary dichotomy on a set of m points *in general position* in $\mathbb{R}^n$; if the points are on the corners of the n-dimensional hypercube (i.e., binary vectors), $m - 1$ nodes are still needed;

- a slightly tighter bound was proven in [16]: only $\lceil 1 + (m-2)/n \rceil$ neurons are needed in the hidden layer for realizing an arbitrary dichotomy on a set of m points which satisfy a more relaxed topological assumption (only the points form a sequence from some subsets are required to be *in general position*); also, the $m - 1$ nodes condition was shown to be the least upper bound needed;

- Arai [2] showed that $m - 1$ hidden neurons are necessary for arbitrary separability (any mapping between input and output for the case of binary-valued units), but improved the bound for the two-category classification problem to $m/3$ (without any condition on the inputs).

These results show that for binary inputs the *size* grows exponentially (as $m \leq 2^n$). Some existence lower bounds for the arbitrary dichotomy problem are (see [14]):

- a *depth*-2 NN requires at least $m / [n \log(m/n)]$ hidden neurons (if $m \geq 3n$);

- a *depth*-3 NN requires at least $2 \, (m/\log m)^{1/2}$ neurons in each of the two hidden layer (if $m \gg n^2$); this bound is identical to the one presented in [20] for $m = 2^n$;

- an arbitrarily interconnected NN without feedback needs $(2m/\log m)^{1/2}$ neurons (if $m \gg n^2$).

Several other bounds for arbitrary Boolean functions can be found in [22]. All of these are: (i) revealing a gap between the upper and the lower bounds, thus encouraging research efforts to reduce (or even close) these gaps; (ii) suggesting that networks with more hidden layers might have a smaller *size*.

---

[1] $\lceil x \rceil$ is the ceiling of x, i.e., the smallest integer greater than or equal to x, and $\lfloor x \rfloor$ is the floor of x, i.e., the largest integer less than or equal to x.
In this paper all the logarithms are taken to base 2.

4. Entropy Bounds

A different approach to classification problems has been presented in [6]; it is based on computing the entropy (see also [1] and [25]) of the data-set. Two bounds on the *number-of-bits* have been proven: a constructive one

$$\#bits \; < \; mn\left\{\lceil \log(D/d)\rceil + 5/2\right\} = O\,(mn) \tag{8}$$

and an existence one $\#bits = O\,(m\log m)$. For arbitrary Boolean functions, the two bound are equivalent as $m = 2^{n}$, which gives $\#bits = O\,(n \cdot 2^{n})$ for both of them. Because we shall build on the $O\,(mn)$ constructive bound (8), we present here a sketch of proof (for details see [6]).

Sketch of proof. Find the examples (from the two different classes) which are the closest to one another: $x_d, y_{d'}$ (the distance between them is d). Translate the origin of the axes into x_d and rotate the axes such as the origin (i.e., x_d) and $y_{d'}$ represent the opposite corners of a hypercube. The side length is $l = d/\sqrt{n}$, and we can use it as a step to quantize the whole space. As there are no examples situated at a distance closer than d, it is certain that there will be no hypercube containing examples from the two different classes.

Because the largest distance is D, there is a ball in $\mathbb{R}^{n}$ of radius D which contains all the m examples (see Fig. 1). All the m_{+} (resp. m_{-}) examples are contained inside a ball of radius D centered in that example from the opposite class which is used to determine the largest distance D: $y_{D'}$ (resp. x_D). There are only three possible cases:

- all the examples are in the sub-space determined by the intersection of two balls (see Fig. 1.a.); any of the balls will now contain all the m examples;
- there are positive examples situated on the other side (with respect to x_D) of the hyperplane orthogonal to $x_D - y_{D'}$ and containing $y_{D'}$ (see Fig. 1.b.); select the farthest positive example and use it as the center of a third ball of radius D; the intersections of the three balls determine the sub-spaces where there are examples, but all of them are inside the ball centered in $y_{D'}$;
- there are negative examples situated on the other side (with respect to $y_{D'}$) of the hyperplane orthogonal to $x_D - y_{D'}$ and containing x_D (see Fig. 1.c.); now all the m examples are inside the ball centered in x_D.
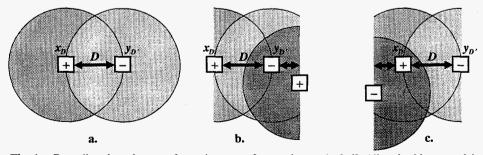

Fig. 1. Bounding the sub-space for a given set of examples: **a.** the balls (discs in this example) determined by the two examples situated at distance D; **b.** modification produced by the farthest positive example (see text); **c.** modification produced by the farthest negative example (see text).

The *number-of-bits* for *one example* can now be computed as $\lceil \log(V_{ball}/v_{hc}) \rceil$ where the volume of the ball of radius D is $V_{ball}(D, n) = \alpha(n) \cdot D^n$, and the volume of the hypercube is $v_{hc}(d, n) = (d/\sqrt{n})^n$:

$$\#bits_{example} = \left\lceil \log\left\{\frac{V_{ball}(D, n)}{v_{hc}(d, n)}\right\} \right\rceil = \left\lceil \log\left\{\frac{\pi^{n/2}D^n}{\Gamma(n/2+1)} \cdot \frac{n^{n/2}}{d^n}\right\} \right\rceil$$

$$< n\left\{\lceil \log(D/d) \rceil + 5/2\right\}. \tag{9}$$

By multiplying with m, the proof is concluded. $\square$

This result has been obtained using a whole ball $V_{ball}(D, n)$ of radius D for upper bounding the space containing the examples. Looking again at Fig. 1 one can see that:

- for the case detailed in Fig. 1.a. all the m examples are inside the intersection of two balls of radius D;
- for the case detailed in Fig. 1.b. all the negative examples m_- are inside a sub-space of the intersection of two balls of radius D;
- for the case detailed in Fig. 1.c. all the positive examples m_+ are inside a sub-space of the intersection of two balls of radius D.

It becomes clear that a better bound on the *number-of-bits* can be obtained if instead of the volume V_{ball} used in (9), one uses the volume of the intersection of two balls $V(D, n)$ as in Fig1.a (which is the worst case).

Proposition 1. The measure (i.e. the 'volume') of the intersection of two balls in $\mathbb{R}^n$ of the same radius $r \in \mathbb{R}^+ \setminus \{0\}$, placed such that the center of each one is on the boundary of the other one, is $V(r, n) = 2\alpha(n-1) r^n \cdot a(n)$ with:

$$a(n) = \frac{n-1}{n} \cdot a(n-2) - \frac{3^{(n-1)/2}}{n \cdot 2^n}.$$

Proof. The given problem can be rewritten as follows: let $c_1 = (0, 0, \dots, 0)$ and $c_2 = (r, 0, \dots, 0)$; compute $V(r, n) = \mu_n(B_n[c_1, r] \cap B_n[c_2, r])$. As it will be seen, $V(r, n)$ can be written as $V(r, n) = k(n) \cdot r^n$ for a suitable *'constant'* $k(n)$. This can be deduced from the observation that, under a homothetic expansion of factor r, $B_n[0, r] \cap B_n[(r, 0, \dots, 0), r]$ is the image of $B_n[0, 1] \cap B_n[(1, 0, \dots, 0), 1]$. Formally, if we define:

$$h_r : \mathbb{R}^n \to \mathbb{R}^n : x \to r \cdot x \tag{10}$$

then

$$\begin{aligned}
V(r, n) &= \mu_n(B_n[0, r] \cap B_n[(r, 0, \dots, 0), r]) \\
&= \mu_n\{h_r(B_n[0, 1] \cap B_n[(1, 0, \dots, 0), 1])\} \\
&= r^n \cdot \mu_n(B_n[0, 1] \cap B_n[(1, 0, \dots, 0), 1]) \\
&= r^n \cdot V(1, n) \tag{11}
\end{aligned}$$

and hence $k(n) = V(1, n)$.

The "geometric" idea for computing $V(r, n)$ goes as follows: assume that $n = 3$ and observe that by symmetry $V(r, n)$ is twice the volume of a segment of a sphere (the part of a sphere situated on one side of a plane). The plane in $n = 3$ becomes a hyperplane in $\mathbb{R}^n$; it is the hyperplane containing the $n - 1$ ball (i.e., from $\mathbb{R}^{n-1}$) representing the 'intersection' of the two balls [5]. The volume of a segment of a sphere can be computed by summing the *volume* of very thin cylinders, or, at the limit, by summing the *area* of the thin discs.

$$V(r, n) \quad = \mu_n \left(B_n [0, r] \cap B_n [(r, 0, \ldots, 0), r] \right) \qquad \text{(due to symmetry)}$$

$$= 2 \mu_n \left(B_n [0, r] \cap B_n [(r, 0, \ldots, 0), r] \cap \{ x \in \mathbb{R}^n : x_1 \geq r/2 \} \right) \text{(Fubini's theorem)}$$

$$= 2 \cdot \int_{r/2}^{r} d\rho \, \mu_{n-1} \left(B_n [0, r] \cap B_n [(r, 0, \ldots, 0), r] \cap \{ x \in \mathbb{R}^n : x_1 = \rho \} \right)$$

$$= 2 \cdot \int_{r/2}^{r} d\rho \, \mu_{n-1} \left(B_n [0, r] \cap \{ x \in \mathbb{R}^n : x_1 = \rho \} \right)$$

$$= 2 \cdot \int_{r/2}^{r} d\rho \, \mu_{n-1} \left(B_{n-1} [0, \sqrt{r^2 - \rho^2}] \right)$$

$$= 2 \cdot \int_{r/2}^{r} d\rho \, (r^2 - \rho^2)^{(n-1)/2} \, \alpha(n-1)$$

$$= 2 \, \alpha(n-1) \, r^{n-1} \cdot \int_{r/2}^{r} d\rho \, [1 - (\rho/r)^2]^{(n-1)/2} \qquad \begin{array}{l} \text{(change of variable:} \\ \sin\theta = \rho/r, \ d\rho = r \cos\theta \, d\theta) \end{array}$$

$$= 2 \, \alpha(n-1) \, r^n \cdot \int_{\pi/6}^{\pi/2} (\cos\theta)^n \, d\theta \qquad\qquad (12.a)$$

$$= 2 \, \alpha(n-1) \, r^n \cdot a(n) \qquad\qquad\qquad (12)$$

where $a(n) = \displaystyle\int_{\pi/6}^{\pi/2} (\cos\theta)^n \, d\theta$.

Using integration by parts we can derive a 'recurrence formula' for $a(n)$. Let:

$$f(\theta) = (\cos\theta)^{n-1} \qquad\quad f'(\theta) = (n-1)(\cos\theta)^{n-2}(-\sin\theta) \qquad (13.a)$$

$$g'(\theta) = \cos\theta \qquad\qquad g(\theta) = \sin\theta \qquad\qquad\qquad\qquad (13.b)$$

we find:

$$a(n) = \left[(\cos\theta)^{n-1} \cdot \sin\theta \right]_{\pi/6}^{\pi/2} - \int_{\pi/6}^{\pi/2} (n-1)(-\sin^2\theta)(\cos\theta)^{n-2} d\theta$$

$$= 0 - \left(\frac{\sqrt{3}}{2}\right)^{n-1} \cdot \frac{1}{2} + (n-1) \cdot \int_{\pi/6}^{\pi/2} (1 - \cos^2\theta)(\cos\theta)^{n-2} d\theta$$

$$= -\frac{1}{2} \cdot \left(\frac{\sqrt{3}}{2}\right)^{n-1} + (n-1) \cdot [a(n-2) - a(n)] \qquad (14)$$

such that:

$$a(n) = \frac{n-1}{n} \cdot a(n-2) - \frac{3^{(n-1)/2}}{n \cdot 2^n} \qquad (15)$$

which concludes the proof. $\square$

For initiating the recurrence we need $a(2)$ and $a(3)$, which are:

$$a(2) = \int_{\pi/6}^{\pi/2} \cos^2\theta \, d\theta = \int_{\pi/6}^{\pi/2} \frac{1 + \cos 2\theta}{2} d\theta = \frac{\pi}{3} - \frac{\sqrt{3}}{8} \qquad (16)$$

$$a(3) = \int_{\pi/6}^{\pi/2} \cos^3\theta \, d\theta = \int_{\pi/6}^{\pi/2} (\cos\theta - \cos\theta \cdot \sin^2\theta) \, d\theta = \frac{5}{24}. \qquad (17)$$

This result can be used to prove a slightly tighter bound on the *number-of-bits*.

Proposition 2. The dichotomy of $m = m_+ + m_-$ examples from $\mathbb{R}^n$ can always be solved with:

$$\#bits < \frac{mn\left\{ \lceil \log(D/d) \rceil + 2 \right\}}{2}. \qquad (18)$$

Proof. We compute the *number-of-bits* required by one example as:

$$\#bits_{example} = \left\lceil \log\left\{ \frac{V(D,n)}{v_{hc}(d,n)} \right\} \right\rceil = \left\lceil \log\left\{ \frac{2\pi^{(n-1)/2} D^n a(n) n^{n/2}}{d^n \Gamma[(n-1)/2 + 1]} \right\} \right\rceil \qquad (19)$$

From (12.a) $\theta \in [\pi/6, \pi/2]$, thus $\cos\theta \le \cos(\pi/6) = \sqrt{3}/2$ leading to:

$$a(n) < \int_{\pi/6}^{\pi/2} (\sqrt{3}/2)^n \, d\theta = \left(\frac{\sqrt{3}}{2}\right)^n \cdot \frac{\pi}{3} \qquad (20)$$

and by using Stirling's formula $n! > \sqrt{2\pi n} \cdot (n/e)^n$ we have:

$$
\begin{aligned}
\#bits_{example} &= \left\lceil 1 + \frac{n}{2}\log\pi - \frac{1}{2}\log\pi + n\log(D/d) + \log\pi - \log3 + \frac{n}{2}\log3 - n + \frac{n}{2}\log n \right. \\
&\qquad \left. - \frac{1}{2}\log\pi - \frac{1}{2}\log(n-1) - \frac{n-1}{2}\log(n-1) + \frac{n-1}{2} + \frac{n-1}{2}\log e \right\rceil \\
&< \lceil\, n\log(D/d) + 1.8396\,n\,\rceil \\
&< n\left\{\lceil\log(D/d)\rceil + 2\right\}.
\end{aligned}
\tag{21}
$$

Because a dichotomy can be solved either for the positive, or for the negative examples, we can choose the smallest of these two sets: $\min(m_+, m_-) \leq m/2$, and the proof is concluded.   □

Finally, from these results we can obtain the following bounds on the *size* of NNs solving a dichotomy problem.

Proposition 3.   The dichotomy of $m = m_+ + m_-$ examples from $\mathbb{R}^n$ can always be solved by a network of neurons with fan-in $= \Delta$ having:

$$
\frac{mn\left\{\lceil\log(D/d)\rceil + 2\right\}}{\Delta^2\log\Delta} < size < \frac{mn\left\{\lceil\log(D/d)\rceil + 2\right\}}{\Delta^2}.
\tag{22}
$$

Proof. It is known [21, 23] that the *weights* of integer representations of Boolean linear threshold functions having Δ inputs can be bounded as:

$$
2^{(\Delta-1)/2} < weight < \frac{(\Delta+1)^{(\Delta+1)/2}}{2^\Delta}
\tag{23}
$$

and by taking logarithms we can obtain the *number-of-bits* required by one *weight* as:

$$
\left\lceil\frac{\Delta}{2}\right\rceil < \#bits_{weight} < \left\lceil\frac{\Delta\log\Delta}{2}\right\rceil.
\tag{24}
$$

The *number-of-bits* of the NN will be:

$$
size\,\Delta\cdot\left\lceil\frac{\Delta}{2}\right\rceil < \#bits_{NN} < size\,\Delta\cdot\left\lceil\frac{\Delta\log\Delta}{2}\right\rceil
\tag{25}
$$

which shows that:

$$
\frac{2\cdot\#bits_{NN}}{\Delta^2} > size > \frac{2\cdot\#bits_{NN}}{\Delta^2\log\Delta}.
\tag{26}
$$

Because the entropy of the data-set is encoded by the NN, there are solutions having $\#bits \leq \#bits_{NN}$; the solution(s) with a 'good' encoding will have $\#bits \cong \#bits_{NN}$ (there are cases [9, 10, 18] for which $\#bits > \#bits_{NN}$), and using (18) the result follows. □

Several remarks are in order:

- the upper bound is even tighter as, from [21], $weigth > 1.618^{\Delta}$ which gives

$$size < \frac{0.72\,mn\left\{\lceil\log(D/d)\rceil + 2\right\}}{\Delta^2}; \tag{27}$$

- the known bounds do not impose any limitation on the *fan-in*, thus in principle one could take $\Delta = n$ which gives

$$\frac{m\left\{\lceil\log(D/d)\rceil + 2\right\}}{n\,\log n} < size < \frac{0.72\,m\left\{\lceil\log(D/d)\rceil + 2\right\}}{n}; \tag{28}$$

- for arbitrary Boolean functions $\lceil\log(D/d)\rceil = \log\sqrt{n}$, giving

$$\frac{m}{2n} < size < \frac{0.36\,m\,\log n}{n} \tag{29}$$

and showing that the bounds are tight (see Section 3 for known the bounds).

All these results can easily be extended to Boolean circuits by taking an encoding of one bit per input.

4. Conclusions

Based on the entropy of the data-set, this paper has presented a new nonconstructive proof on the *size* of NNs for solving a dichotomy. The resulting lower and upper bounds are tight.

Although the proof for the *number-of-bits* is constructive, the shape of the bounding space does not lend itself easily to practical applications. Bounding the space with a ball, or the intersection of two balls—which, as we have seen, is theoretically possible—is computationally too difficult. For all practical cases, the simplest bounding space is a hypercube. By taking a hypercube of side length $2D$ the problem can be solved, but unfortunately we have to pay by a logarithmic increase on the *number-of-bits*:

$$m\left\lceil\log\left\{\frac{(2D)^n}{(d/\sqrt{n})^n}\right\}\right\rceil < m\left[5n/2 + n\log(D/d) + (n\log n)/2\right] = O\,(mn\log n).$$

Presently, we are working on this particular aspect by trying to use other co-ordinates (e.g., polar co-ordinates instead of the rectangular ones).

References

1. Abu-Mostafa, Y.S.: Connectivity versus entropy, in D.Z. Anderson (ed.): *Neural information processing systems* (NIPS*87, Denver, CO), Amer. Inst. Physics, NY (1988) 1–8.
2. Arai, M.: Bounds on the number of hidden units in binary-valued three-layer neural networks, *Neural Networks* 6(6) (1993) 855–860.
3. Arbib, M.A. (ed.): *The handbook of brain theory and neural networks*, MIT Press, Cambridge, MA (1995).

4. Baum, E.B.: On the Capabilities of Multilayer Perceptrons, *J. of Complexity* **4** (1988) 193–215.

5. Baum, E.B.: When are *k*–nearest neighbour and back propagation accurate for feasible size sets of examples ?, draft in E. Amaldi and E. Mayoraz: *Mathematical foundations of artificial neural networks* (summer school lecture notes), Swiss Federal Institute of Technology Lausanne and Kurt Bösch Academic Institute Sion, September 1992. In S.J. Hanson, G.A. Drastal and R.L. Rivest (eds.): *Computational learning theory and natural learning systems—Vol. I: Constraints and prospects*, Bradford/MIT Press, 1994.

6. Beiu, V.: Entropy bounds for classification algorithms, *Neural Network World* **6**(4) (1996) 497–505.

7. Beiu, V.: Digital integrated circuit implementations, *Chapter E1.4* in [13], E1.4:1–34.

8. Beiu, V.: *VLSI complexity of discrete neural networks*, Gordon and Breach, Newark (1997).

9. Beiu, V., Taylor, J.G.: VLSI optimal neural network learning algorithm, in D.W. Pearson, N.C. Steele and R.F. Albrecht (eds.): *Artificial neural nets and genetic algorithms* (ICANNGA'95, Alès, France), Springer-Verlag, Vienna (1995) 61–64.

10. Beiu, V., Taylor, J.G.: Direct synthesis of neural networks, *Proc. MicroNeuro'96* (Lausanne, Switzerland), IEEE CS Press, Los Alamitos, CA (1996) 257–264.

11. Bishop, C.M.: *Neural networks for pattern recognition*, Oxford University Press, Oxford, UK (1995).

12. Bruck, J., Smolensky, R.: Polynomial threshold functions, AC^0 functions and spectral norms, *Res. Rep. RJ 7410* (67387), 11/15/89, IBM Yorktown Heights, NY (1989). Also in *SIAM J. Computing* **21**(1) (1992) 33–42.

13. Fiesler, E., Beale, R. (eds.): *Handbook of neural computation.* Oxford Univ. Press and the Inst. of Physics Publishing, NY (1996).

14. Hassoun, M.H.: *Fundamentals of artificial neural networks*, MIT Press, Cambridge, MA (1995).

15. Hecht-Nielsen, R., *Neurocomputing*, Addison-Wesley, Reading, MA (1989).

16. Huang, S.-C., Huang, Y.-F.: Bounds on the number of hidden neurons of multilayer perceptrons in classification and recognition, *IEEE Trans. on Neural Networks* **2**(1) (1991) 47–55.

17. Kohonen, T.: *Self-organizing and associative memory*, Springer-Verlag, Berlin (1988).

18. Krishnamoorthy, A.V., Paturi, R., Blume, M., Linden, G.D., Linden L.H., Esener, S.C.: Hardware tradeoffs for Boolean concept learning, in *Proc. World Conf. on Neural Networks '94* (WCNN'94, San Diego), Lawrence Erlbaum & INNS Press, Hillsdale (1994) vol. 1, 551–559.

19. Lupanov, O.B.: The synthesis of circuits from threshold elements, *Problemy Kibernetiki* **20** (1973) 109–140.

20. Neciporuk, E.I.: The synthesis of networks from threshold elements, *Problemy Kibernetiki* **11** (1964) 49–62. English translation in *Automation Express* **7**(1) 35–39 and **7**(2) 27–32.

21. Parberry, I.: *Circuit complexity and neural networks*, MIT Press, Cambridge, MA (1994).

22. Paugam-Moisy, H.: *Optimisation des réseaux des neurones artificiels.* PhD thesis, Laboratoire de l'Informatique du Parallélisme LIP-IMAG, École Normale Supérieure de Lyon (46 Allée d'Italie, 69364 Lyon, France) (1992).

23. Raghavan, P.: Learning in threshold networks: a computational model and applications. *Tech. Rep. RC 13859*, IBM Res. (1988). Also in *Proc. 1st Workshop on Computational Learning Theory* (Cambridge, MA), ACM Press (1988) 19–27.

24. Siu, K.-Y., Roychowdhury, V., Kailath, T.: Depth-size tradeoffs for neural computations, *IEEE Trans. on Comp.* **40**(12) (1991) 1402–1412.

25. Williamson, R.C.: ε-entropy and the complexity of feedforward neural networks, in R.P. Lippmann, J.E. Moody and D.S. Touretzky (eds.): *Neural information processing systems* (NIPS*90, Denver, CO), Morgan Kaufmann, San Mateo (1991) 946–952.

CONTRIBUTIONS MUST BE SENT TO:

Prof. José Mira
Dpto. Inteligencia Artificial, UNED
Senda del Rey, s/n
E - 28040 MADRID, Spain
E-mail: iwann97@dia.uned.es

Phone: + 34 1 3987155
Fax: + 34 1 3986697

IMPORTANT DATES

Final Date for Submission January 15, 1997
Notification of Acceptance March 1997
Work-Conference June 4-6, 1997

INSCRIPTION, TRAVEL AND HOTEL INFORMATION

ULTRAMAR EXPRESS
Diputació, 238, 3°
E-08007 BARCELONA, Spain

Phone: +34 3 4827140
Fax: +34 3 4827158

IBERIA and AVIACO will be the official carriers for IWANN'97, offering special rates and conditions. International code for special rate: BT71B21MPE0038.

POSSIBILITY OF GRANTS

The Organization Committee of IWANN'97 will provide a very limited number of full or partial grants. Please contact the WWW address for further information.
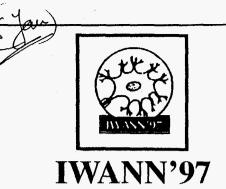
STEERING COMMITTEE

- Joan Cabestany , Universidad Politécnica de Catalunya (E)
- José Mira Mira, UNED (E)
- Alberto Prieto, Universidad de Granada (E)
- Francisco Sandoval, Universidad de Malaga (E)

ORGANIZATION COMMITTEE

- *Joan Cabestany and Francisco Sandoval (E), Co-chairmen*
- Michael Arbit, University of Southern California (USA)
- Senén Barro, Universidad de Santiago (E)
- Gabriel de Blasio, Univ. de Las Palmas de Gran Canaria (E)
- Trevor Clarkson, King's College London (UK)
- Ana Delgado, UNED (E)
- Dante Del Corso, Politecnico de Torino (I)
- Belén Esteban-Sánchez, ITC (E)
- Tamas D. Gedeon, University of New South Wales (AUS)
- Karl Goser, Universität Dortmund (G)
- Jeanny Herault, Institute National Polytechnique de Grenoble (F)
- Jaap Hoekstra, Delft University of Technology (NL)
- Shunsuke Sato, Osaka University (Jp)
- Igor Shevelev, Russian Academy of Science(R)
- Cloe Taddei-Ferretti, Istituto di Cibernetica, CNR (I)
- Marley Vellasco, Pontificia Universidade Catolica do Rio de Janeiro (Br)
- Michel Verleysen, Universite Catholique de Louvain-la-Neuve (B)

PROGRAMME COMMITTEE

- *José Mira and Alberto Prieto, Co-chairmen (E)*
- Igor Aleksander, Imperial Coll. of Science Technology and Medicine (UK)
- José Ramón Alvarez, UNED (E)
- Shun-Ichi Amari, University of Tokyo (Jp)
- Xavier Arreguit, CSEM (CH)
- François Blayo, Univ. Paris 1 (F)
- Leon Chua, University of California (USA)
- Marie Cottrell, Univ. Paris 1 (F)
- Akira Date, Tokyo University of Agriculture and Technology (Jp)
- Antonio Díaz-Estrella, Universidad de Málaga (E)
- M. Duranton, Phillips (F)
- Reinhard Eckhorn, Philips University (D)
- Kunihiko Fukushima, Osaka University (Jp)
- Patrik Garda, Univ. Pierre et Marie Curie (F)
- Anne Guérin-Dugué, INPG (F)
- Martin Hasler, EPFL (CH)
- Mohamad H . Hassoun, Wayne State University (USA)
- Gonzalo Joya, Universidad de Málaga (E)
- Simon Jones, JERI Loughborough Univ. of Tech. (UK)
- Christian Jutten, INPG (F)
- H. Klar, Technische Universität Berlin (G)
- K.Nicholas Leibovic, Univ. Buffalo (USA)
- J.Lettvin, MIT (USA)
- Francisco Javier López Aligué, Universidad de Extremadura (E)
- Jordi Madrenas, UPC (E)
- Pierre Marchal, CSEM (CH)
- Juan Manuel Moreno, UPC (E)
- Josef A. Nossek, Der Technischen Universität München (G)
- Julio Ortega, Universidad de Granada (E)
- Francisco José Pelayo, Universidad de Granada (E)
- Franz Pichler, Johannes Kepler Universität Linz (A)
- Vicenzo Piuri, Politecnico di Milano (I)
- Leonardo Reyneri, Politecnico di Torino (I)
- Tamas Roska, Hungarian Academy of Sciences (H)
- E. Sanchez-Sinencio, Texas A&M Univ. (USA)
- J. Simoes Da Fonseca, Faculty of Medicine of Lisbon (P)
- John G. Taylor, King's College London (UK)
- Carme Torras, Instituto de Cibernética del CSIC-UPC (E)
- Philip Treleaven, University College London (UK)
- Elena Valderrama, Centro Nacional de Microelectrónica (E)

IWANN'97

INTERNATIONAL WORK-CONFERENCE ON ARTIFICIAL AND NATURAL NEURAL NETWORKS

Biological and Artificial Architectures, Technologies and Applications

Final Call for Papers

Lanzarote - Canary Islands, Spain
June 4-6, 1997

ORGANIZED BY

Universidad Nacional de Educación a Distancia (UNED), Madrid
Universidad de Las Palmas de Gran Canaria
Universidad Politécnica de Catalunya
Universidad de Málaga
Universidad de Granada

IN COOPERATION WITH

Asociación Española de Redes Neuronales (AERN)
IFIP Working Group in Neural Computer Systems, WG10.6
Spanish RIG IEEE Neural Networks Council
UK&RI Communication Chapter of IEEE

Contact URL http://petrus.upc.es/iwann97.html

IWANN'97. The fourth International Workshop on Artificial Neural Networks, now changed to International Work-Conference on Artificial and Natural Neural Networks, will take place in Lanzarote, Canary Islands (Spain) from 4 to 6 of June, 1997. This biennial meeting with focus on biologically inspired and more realistic models of natural neurons and neural nets and new hybrid computing paradigms, was first held in Granada (1991), Sitges (1993) and Torremolinos, Malaga (1995) with a growing number of participants from more than 20 countries and with high quality papers published by Springer-Verlag (LNCS 540, 686 and 930).

SCOPE

Neural computation is considered here in the dual perspective of analysis (as science) and synthesis (as engineering). As a science of analysis, r   l computation seeks to help neurology, brain theory, and cognitive psychology in the understanding of the functioning of the Nervous Systems by means of computational models of neurons, neural nets and subcellular processes, with the possibility of using electronics and computers as a «laboratory» in which cognitive processes can be simulated and hypothesis proven without having to act directly upon living beings.

As a synthesis engineering, neural computation seeks to complement the symbolic perspective of Artificial Intelligence (AI), using the biologically inspired models of distributed, self-programming and self-organizing networks, to solve those non-algorithmic problems of function approximation and pattern classification having to do with changing and only partially known environments. Fault tolerance and dynamic reconfiguration are other basic advantages of neural nets.

In the sea of meetings, congresses and workshops on ANN's, IWANN'97 focus on the three subjects that most concern us:

(1) The seeking of biologically inspired new models of local computation architectures and learning along with the organizational principles behind of the complexity of intelligent behavior.

(2) The searching for some methodological contributions in the analysis and design of knowledge-based ANN's, instead of «   l nets», and in the reduction of the knowledge level to the sub-symbolic implementation level.

(3) The cooperation with symbolic AI, with the integration of connectionist and symbolic processing in hybrid and multi-strategy approaches for perception, decision and control tasks, as well as for case-based reasoning, concepts formation and learning.

To contribute to the posing and partially solving of these global topics, IWANN'97 offer a brain-storming interdisciplinary forum in advanced Neural Computation for scientists and engineers from biology neuroanatomy, computational neurophysiology, molecular biology, biophysics, linguistics, psychology, mathematics and physics, computer science, artificial intelligence, parallel computing, analog and digital electronics, advanced computer architectures, reverse engineering, cognitive sciences and all the concerned applied domains (sensory systems and signal processing, monitoring, diagnosis, classification and decision making, intelligent control and supervision, perceptual robotics and communication systems).

Contributions on the following and related topics are welcome.

TOPICS

1. *Biological Foundations of Neural Computation:* Principles of brain organization. Neuroanatomy and Neurophysiology of synapses, dendrodendritic contacts, neurons and neural nets in peripheral and central areas. Plasticity, learning and memory in natural neural nets. Models of development and evolution. The computational perspective in Neuroscience.

2. *Formal Tools and Computational Models of Neurons and Neural Nets Architectures:* Analytic and logic models. Object oriented formulations. Hybrid knowledge representation and inference tools (rules and frames with analytic slots). Probabilistic, bayesian and fuzzy models. Energy related models.

3. *Plasticity Phenomena (Maturing, Learning and Memory):* Biological mechanisms of learning and memory. Computational formulations using correlational, reinforcement and minimization strategies. Conditioned reflex and associative mechanisms. Inductive- deductive and abductive symbolic-subsymbolic formulations. Generalization.

4. *Complex Systems Dynamics:* Self-organization, cooperative processes, autopoiesis, emergent computation, synergetic, evolutive optimization and genetic algorithms. Self-reproducing nets. Self-organizing feature maps. Simulated evolution. Social organization phenomena.

5. *Cognitive Science and AI:* Hybrid knowledge based system. Neural networks for knowledge modeling, acquisition and refinement. Natural language understanding. Concepts formation. Spatial and temporal planning and scheduling. Intentionality.

6. *Neural Nets Simulation, Emulation and Implementation:* Environments and languages. Parallelization, modularity and autonomy. New hardware implementation strategies (FPGA's, VLSI, neurodevices). Evolutive architectures. Real systems validation and evaluation.

7. *Methodology for Data Analysis, Task Selection and Nets Design.*

8. *Neural Networks for Perception:* Biologically inspired preprocessing. Low level processing, source separation, sensor fusion, segmentation, feature extraction, adaptive filtering, noise reduction, texture, stereo correspondence, motion analysis, speech recognition, artificial vision, and hybrid architectures for multisensorial perception.

9. *Neural Networks for Communications Systems:* Modems and codecs, network management, digital communications.

10. *Neural Networks for Control and Robotics:* Systems identification, motion planning and control, adaptive, predictive and model-based control systems, navigation, real time applications, visuo-motor coordination.

CALL FOR PAPERS

The Programme Committee seeks for original papers on the above mentioned topics. Authors should pay special attention to explanation of theoretical and technical choices involved, point out possible limitations and describe the current state of their work. All received papers will be reviewed by the Programme Committee. Accepted papers may be presented orally or as poster panels, however all accepted contributions will be published in full length (LNCS Springer-Verlag Series).

INSTRUCTIONS TO AUTHORS

Five copies (one original and four copies) of the paper must be submitted. The paper must not exceed 10 pages, including figures, tables and references. It should be written in English on A4 paper, in a Times font 10 point in size, without page numbers (please, indicate the order by numbering the reverse side of the sheets with a pencil) . The printing area should be 12.2 x 19.3 cm. The text should be justified to occupy the full line width, and using one-line spacing. Headings (12 point, bold) should be capitalized and aligned to the left. Title (14 point, bold) should be centered. Abstract and affiliation (9 point) must be also included.

If possible, please make use of the latex/plaintex style file available in the WWW page: http://petrus.upc.es/iwann97.html, where you can get more detailed instructions to the authors. In addition, one sheet must be attached including: Title and authors names, list of five keywords, the Topic the paper fits best, preferred presentation (oral or poster) and the corresponding author (name, postal and e-mail addresses, phone and fax numbers).