Chapter 1

Multilayer Feedforward Networks

Definition 1. *(Essentially bounded). A function u defined almost everywhere with respect to Lebesgue measure v on a measurable set $\Omega \in \mathbb{R}^n$ is said to be* ***essentially bounded*** *on Ω if $|u(x)|$ is bounded almost everywhere on Ω. We denote $u \in L^\infty(\Omega)$ with the norm $\|u\|_{L^\infty(\Omega)} = inf(\lambda|\{x : |u(x)| \geq \lambda\} = 0) = ess\sup_{x\in\Omega}|u(x)|$*

We have that $L^\infty(\mathbb{R})$ is the space of essentially bounded functions.

Examples and counterexamples of functions essentially bounded.

- $f : \Omega \to$

Definition 2. *(Locally essentially bounded). A function u defined almost everywhere with respect to Lebesgue measure on a domain Ω (a domain is an open set in $\mathbb{R}^n$) is said to be* ***locally essentially bounded*** *on Ω if for every compact set $K \subset \Omega$, $u \in L^\infty(K)$. We denote $u \in L^\infty_{loc}(K)$*

Definition 3. *We say that a set of functions $F \subset L^\infty_{loc}(\mathbb{R})$ is dense in $C(\mathbb{R}^n)$ if for every function $g \in C(\mathbb{R}^n)$ and for every compact $K \subset \mathbb{R}^n$, there exist a sequence of functions $f_j \in F$ such that $lim_{j\to\infty}\|g - f_j\|_{L^\infty(K)}$*

Definition 4. *Let M denote the set of functions which are in $L^\infty_{loc}(\mathbb{R})$ and have the following property. The closure of the set of points of discontinuity of any function in M is of zero Lebesgue measure.*

Proposition 1. *This implies that for any $\sigma \in M$, interval $[a, b]$. and $\delta > 0$, there exists a finite number of open intervals, the union of which we denote by U, of measure δ, such that σ is uniformly continuous on $[a, b]/U$.*

Definition 5. *suport*

Definition 6. *(Multilayer feedforward networks) The general architecture of a multilayer feedforward network, MFN, consist of:*

- *input layer: n-input units, x*
- *one/more hidden layers : intermediate processing units*
- *output layer: m output-units $f(x)$*

function that a MFN compute is:

$$f(x) = \sum_{j=1}^{k} \beta_j \cdot \sigma(w_j \cdot x - \theta_j)$$

- $x = (x_1, ..., x_n)$ *input-vector*
- *k: # of processing-units in the hidden layer*
- $w = (w_1, ..., w_n)$*: weights vector*
- $\sigma : \mathbb{R} \to \mathbb{R}$ *activation function*
- θ *treshold value: ???*
- β

We take $C(\mathbb{R}^n)$ to be the family of real world functions that one may wish to approximate with feedforward network architectures

Definition 7. $\mathcal{C}_0^\infty$ *functions* $\mathcal{C}^\infty$ *with compact support.*

Definition 8. *Convergència uniforme)*

Definition 9. $\varphi : I \to \mathbb{R}$ *is uniformly continuous on* I *if* $\forall \epsilon > 0 \exists\ \delta > 0$ *such that* $|\varphi(x) - \varphi(y)| < \epsilon$ *whenever* $|x - y| < \delta$

Theorem 1. *Let* $\sigma \in M$*. Set*

$$\sum_n = span\{\sigma(w \cdot x + \theta) : w \in \mathbb{R}^n, \theta \in \mathbb{R}\}$$

Then $\sum_n$ *is dense in* $\mathcal{C}(\mathbb{R}^n)$ *if and only if* σ *is not an algebraic polynomial.*