When Neurons Fail

El Mahdi El Mhamdi, Rachid Guerraoui

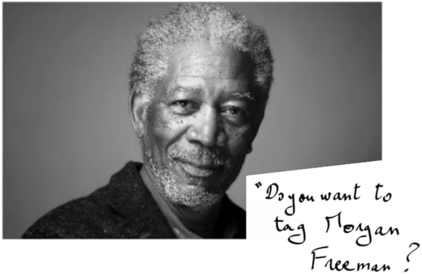
BDA, Chicago
July 25th, 2016

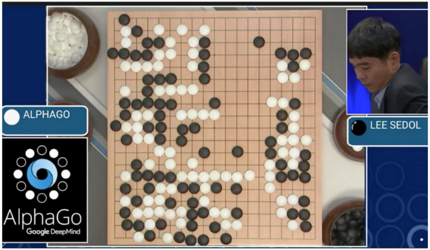
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Laboratoire de Programmation Distribuée
Distributed Programming Laboratory

Table of Contents

NNs everywhere





Model

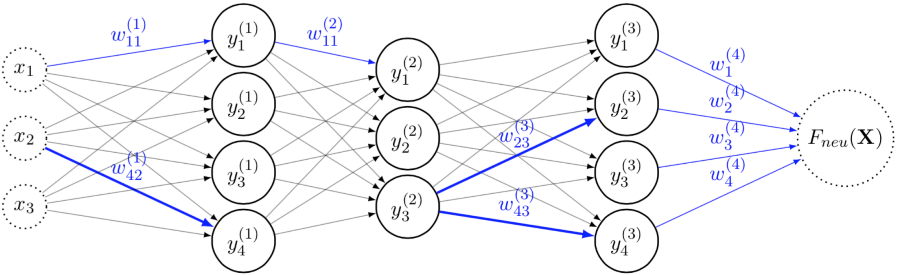


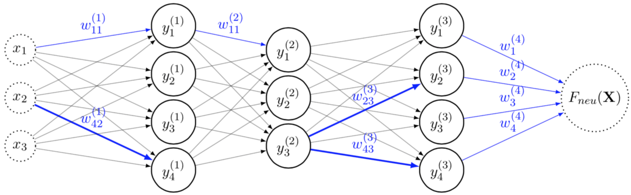
Figure : Feed forward neural network

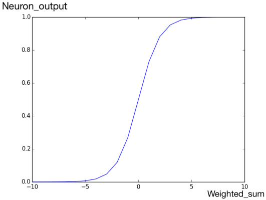
Nodes: neurons
Links: synapses

Model

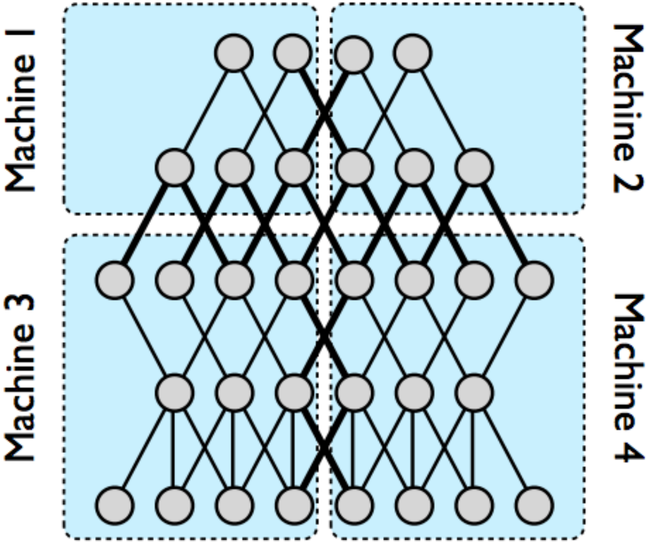
$$F_{neu}(\mathbf{X}) = \sum_{i=1}^{N_L} w_i^{(L+1)} y_i^{(L)}$$

with $y_j^{(l)} = \varphi(s_j^{(l)})$ for $1 \leq l \leq L$; $y_j^{(0)} = x_j$ and $s_j^{(l)} = \sum_{i=1}^{N_{l-1}} w_{ji}^{(l)} y_i^{(l-1)}$

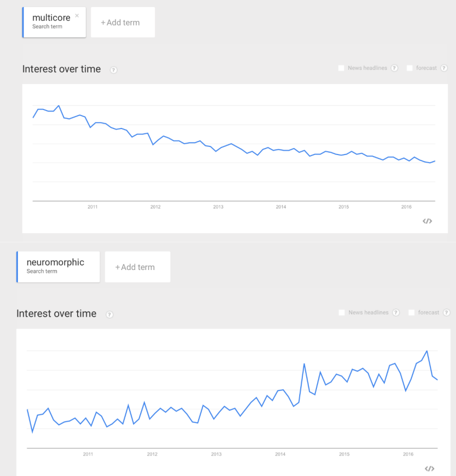




Software simulated NN

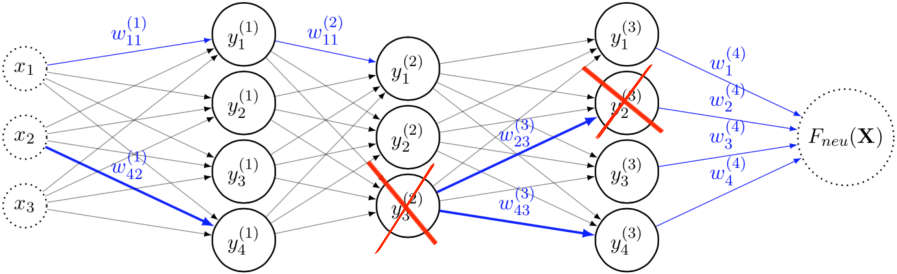


Hardware-based NNs



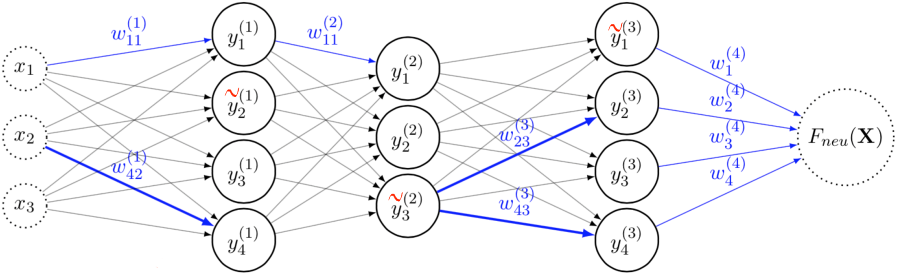
SyNAPSE (DARPA, IBM), Human Brain Project (SP9 on neuromorphic), Brains in Silicon at Stanford...

How robust is this?



Crash failure: a component stops working.

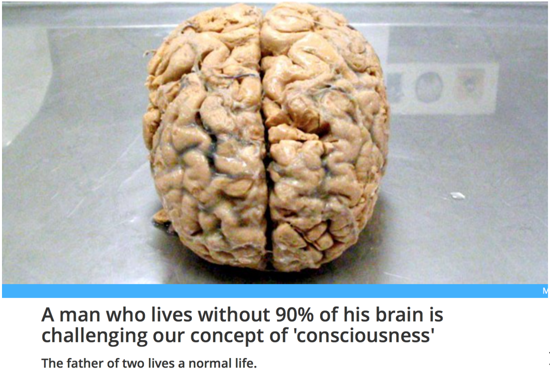
How robust is this?



Byzantine failure: a component sends arbitrary values.

Biological plausibility

Examples of extreme robustness in nature

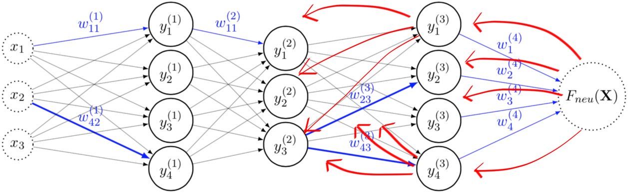


[1]

[1] Feuillet et al., 2007. Brain of a white-collar worker. *Lancet (London, England)*, 370(9583), p.262.

Classical training leads to non-robust NN

E: difference between desired and actual outputs on a training set

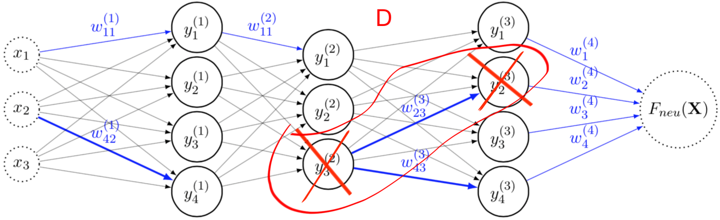


$$\Delta w^{(l)}_{ij} = -\frac{dE}{dw^{(l)}_{ij}}$$

$\exists$ robust weight distribution $\mapsto$ Reach them with learning !

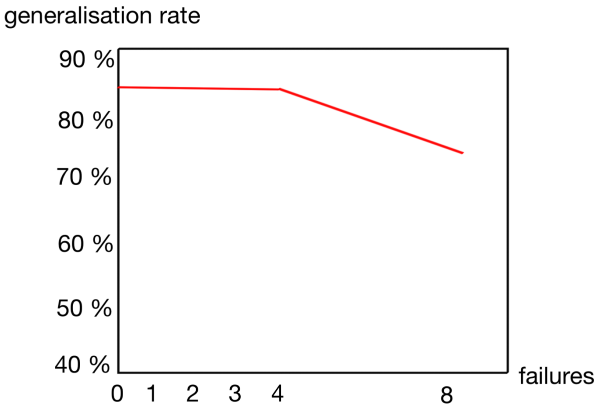
Dropout

Randomly switch neurons off during the training phase
Kerlirzin and Vallet (1991, 1993), Hinton et al. (2012, 2014)



$$\text{Minimize } E_{av} = \sum_D E^D P(D) \text{ where } P(D) = (1-p)^{|D|} p^{(N-|D|)}$$

Experimentally observed robustness



[2]

- Over-provisionning
- Upper-bound ?

[2] from Kerlirzin 1993, edited

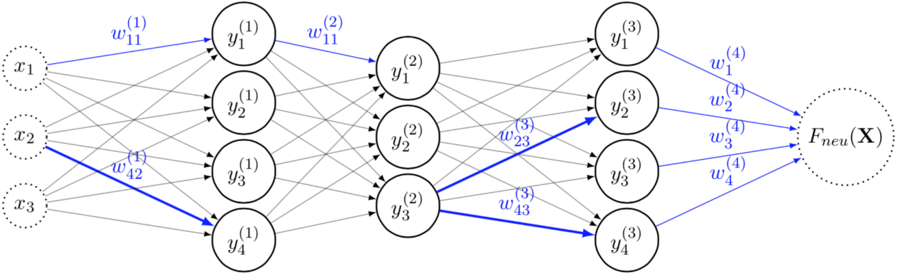
Table of Contents

Given a precision ϵ, derive a tight bound on failures to keep ϵ-precision for a any neural network[3] approximating a function F

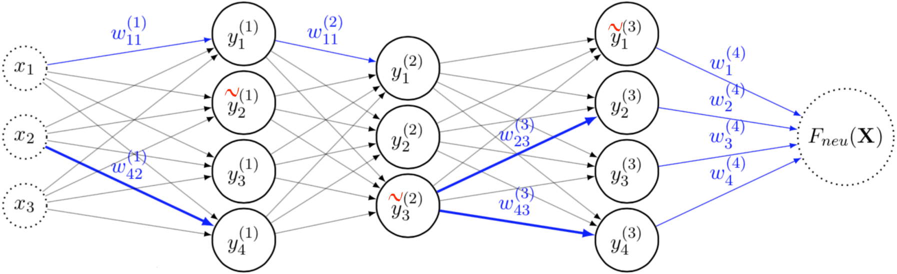
[3] note: learning is taken for granted

Theoretical background: universality

Theorem[4]: $\forall(F, \epsilon)$, $\exists$ NN generating F_{neu} s.t $\|F_{neu} - F\| < \epsilon$



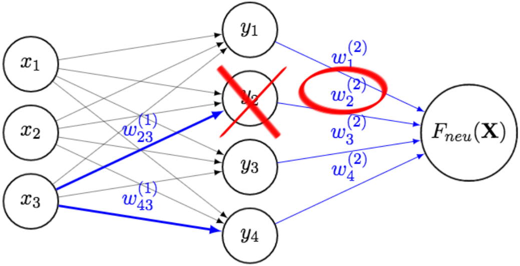
[4]Cybenko 1989, Horkink 1991



- Minimal networks are not robust [5]
- Given over-provision ϵ' ($\epsilon' < \epsilon$), what condition on failures to preserve ϵ-precision?

[5] not to mention: impossible to derive

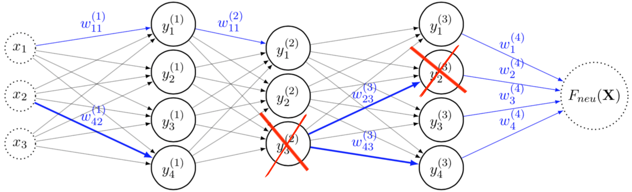
Table of Contents

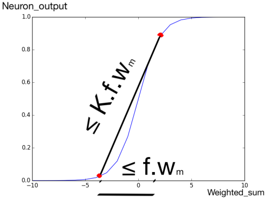


$$f \leq \frac{\epsilon - \epsilon'}{w_m}$$

- More over-provision $\mapsto$ more robustness
- Unequal weight distribution $\mapsto$ single point of failure
- No Byzantine FT $\mapsto$ bounded synaptic capacity

Multilayer networks, Byzantine failures





- Failure at layer l propagates though layers $l' > l$ (Byz and crash).
- Factors: weights, |layers|, |neurons|, Lipschitz coef. of φ
- Total error propagated to the output should be $\leq \epsilon - \epsilon'$

Multilayer networks, Byzantine failures

- Bounded channel capacity (otherwise no robustness to Byzantine)
- Propagated error $\leq C \sum_{l=1}^{L} \left(f_l K^{L-l} w_m^{(L+1)} \prod_{l'=l+1}^{L} (N_{l'} - f_{l'}) w_m^{(l')} \right)$

  C: capacity, K: Lipschitz coeff., $w_m^{(l)}$ maximal weight to layer l
  N_l: $|\text{neurons}|$, f_l: $|\text{failures}|$

How to read the formula

$$C\sum_{l=1}^{L}\left(f_l K^{L-l} w_m^{(L+1)} \prod_{l'=l+1}^{L}(N_{l'}-f_{l'})w_m^{(l')}\right) \leq \epsilon - \epsilon'$$

How to read the formula

$$C\sum_{l=1}^{L}\left(f_l K^{L-l} w_m^{(L+1)} \prod_{l'=l+1}^{L}(N_{l'}-f_{l'})w_m^{(l')}\right) \leq \epsilon - \epsilon'$$

worst-case propagated error

How to read the formula

$$C \sum_{l=1}^{L} \left(f_l K^{L-l} w_m^{(L+1)} \prod_{l'=l+1}^{L} (N_{l'} - f_{l'}) w_m^{(l')} \right) \leq \epsilon - \epsilon'$$

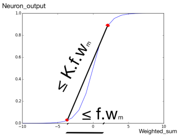
error margin permitted by the over-provision

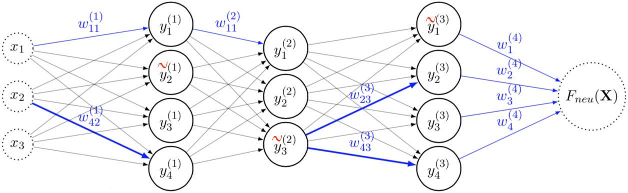
How to read the formula

$$C \sum_{l=1}^{L} \left(f_l K^{L-l} w_m^{(L+1)} \prod_{l'=l+1}^{L} (N_{l'} - f_{l'}) w_m^{(l')} \right) \leq \epsilon - \epsilon'$$

Error (at most C is transmitted) at f_l neurons in layer l propagating through $l' > l$.

$(N_{l'} - f_{l'})$: only correct neurons propagating it, multiplying by $Kw_m^{(l')}$.



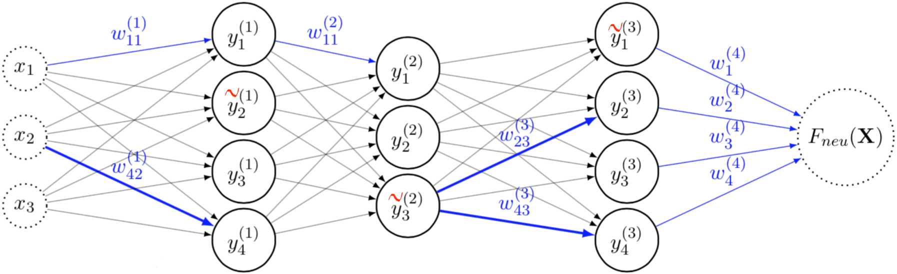


Unbounded capacity

Taking $C \mapsto \infty$

$$C \sum_{l=1}^{L} \left(f_l K^{L-l} w_m^{(L+1)} \prod_{l'=l+1}^{L} (N_{l'} - f_{l'}) w_m^{(l')} \right) \leq \epsilon - \epsilon'$$

Then $\forall\ l\ f_l = 0$
No Byzantine FT.



- Generalization to synaptic failures.
- Applications of the bound (Memory cost, neuron duplication, synchrony)
- Other neural computing models.

Questions ?

More details: `https://infoscience.epfl.ch/record/217561`