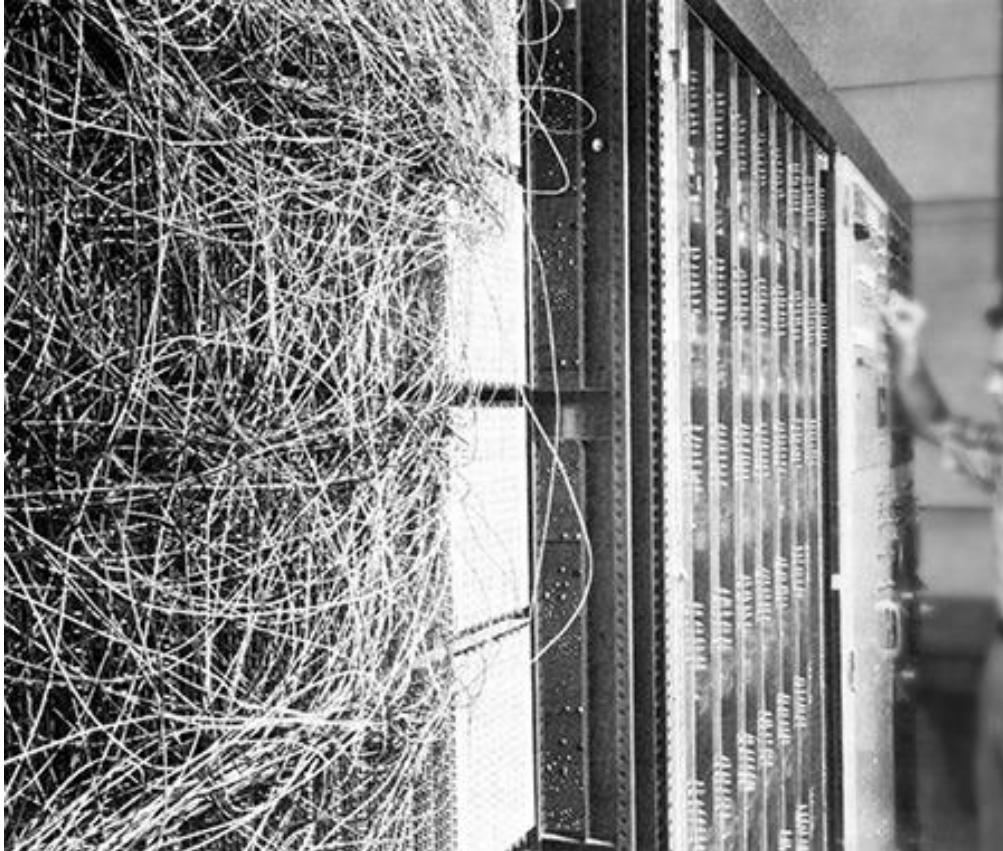


Deep Learning

Taiane Coelho Ramos - 6426955

Perceptron: 50's



1989: Yan LeCun

Reconhecedor de Códigos
Postais

Big Data

A quote from Dan Ariely, “Big data is like teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it ...”

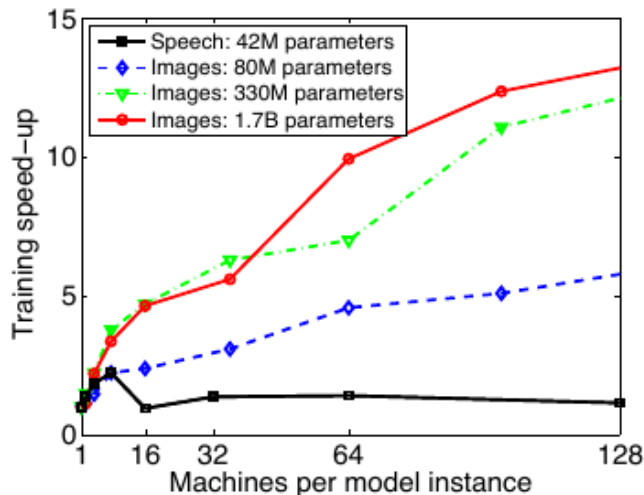


O que é Deep Learning

- Aprender como o cérebro de um bebê
- Algoritmos não-supervisionados
- Extrai características para a classificação

Problemas com Paralelização

“Don’t do it for the sake of parallelization. Strongly consider the programming complexity of approaches if you decide to proceed.”



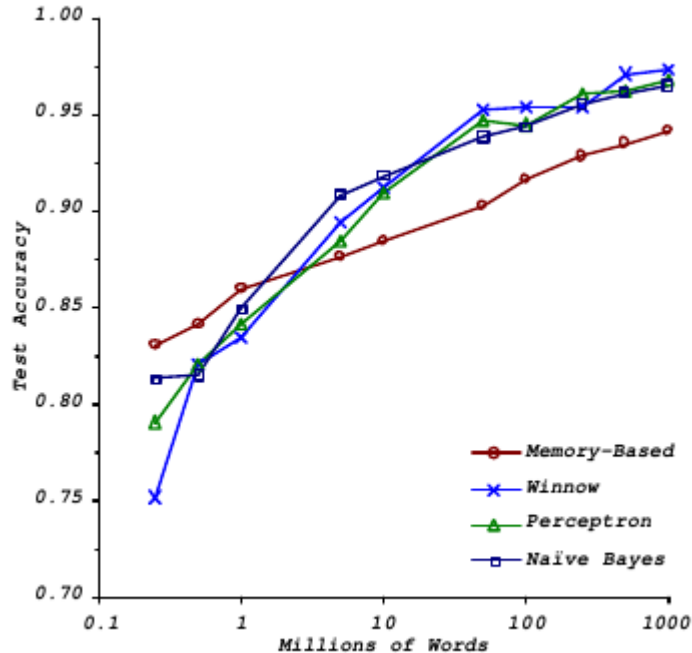
Fonte: Allreduce (MPI) vs. Parameter server approaches - John Langford, 2014

Fonte: Large Scale Distributed Deep Networks - Jeffrey Dean et al., 2012

Problemas com Paralelização

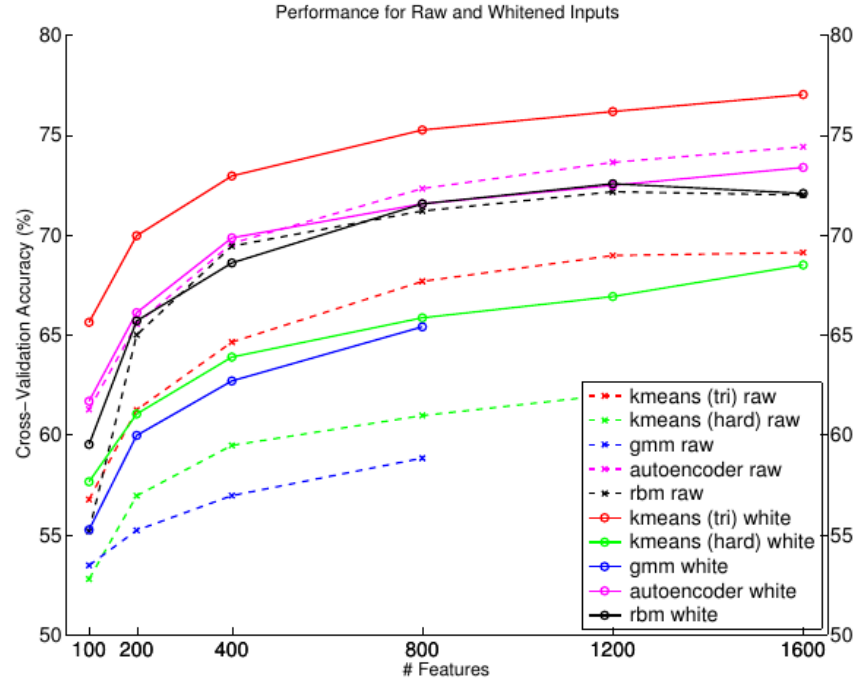
- **Multiplicação de matrizes**
código C: 3m24.843s
MatLab: 4.095059s (BLAS)
- **We want to run mathematical algorithms (classification and clustering) in a complicated architecture (distributed system)**
- **But we are like at the time point before optimized BLAS was developed**

Supervised Learning



“It’s not the best algorithm that wins. Its who has the most data.” - Andrew Ng

Unsupervised Learning



“It’s who has the biggest model that wins.” - Andrew Ng

Precisamos de poder computacional!

Qual método de HPC utilizar?

- **If you are locked into a particular piece of hardware or cluster software, make the best of it (MPI, Hadoop)**
- **If your data can easily be copied onto a single machine, GPU**
- **If your data is of a multimachine scale you must do some form of cluster parallelism.**

Qual método de HPC utilizar?

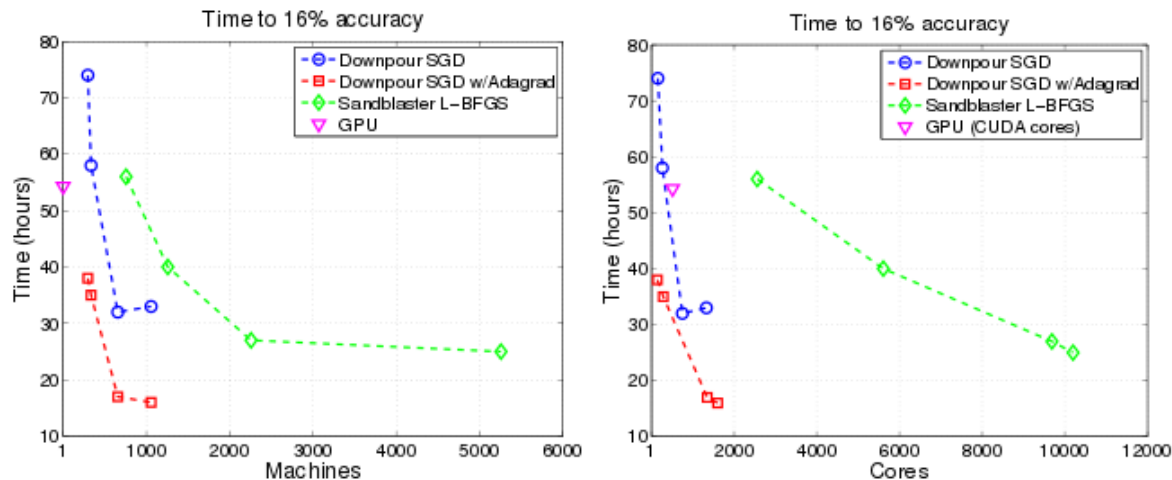


Figure 5: Time to reach a fixed accuracy (16%) for different optimization strategies as a function of number of the machines (left) and cores (right).

Exemplos de Aplicações

Identificação e placas de transito



Performance da Rede: 99,46%

Melhor Humano: 99,22%

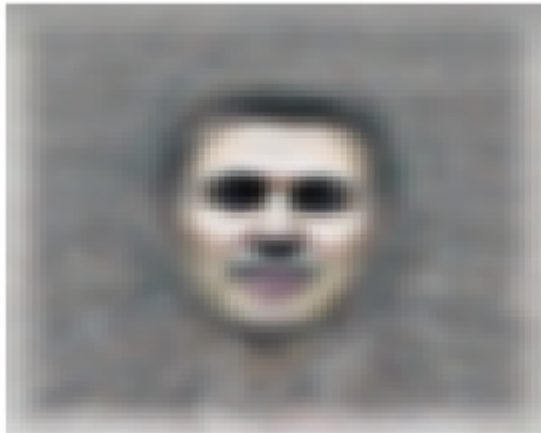
Média Humano: 98,81%

Fig. 3. Random representatives of the 43 traffic sign classes in the GTSRB dataset.

Exemplo de Aplicações

Google Brain

“we would like to understand if it is possible to build a face detector from only unlabeled images.”



1000 máquinas = 16000 cores
3 dias de treinamento

37000 imagens
13026 faces

Deep Net: 81% acerto
Filtro Linear: 74% acerto

Exemplo de Aplicações

Google Brain



Deep Net

gatos: 74.8% acerto

corpos: 76.7% acerto

Filtros Lineares

gatos: 67.2% acerto

corpos: 68.1% acerto

Possíveis aplicações futuras

- Busca pela ideia dentro de um texto
- Tradução usando o “pensamento” de uma frase

THANK YOU

FOR YOUR ATTENTION