

SCM 2020

Prof. Caetano R. Miranda

IFUSP, São Paulo - Brazil

cmiranda@if.usp.br

Camilo A. F. Salvador

IFUSP, São Paulo - Brazil

csalvador@usp.br

APRENDIZADO DE MÁQUINA - LAB 06B
REDES NEURAIS

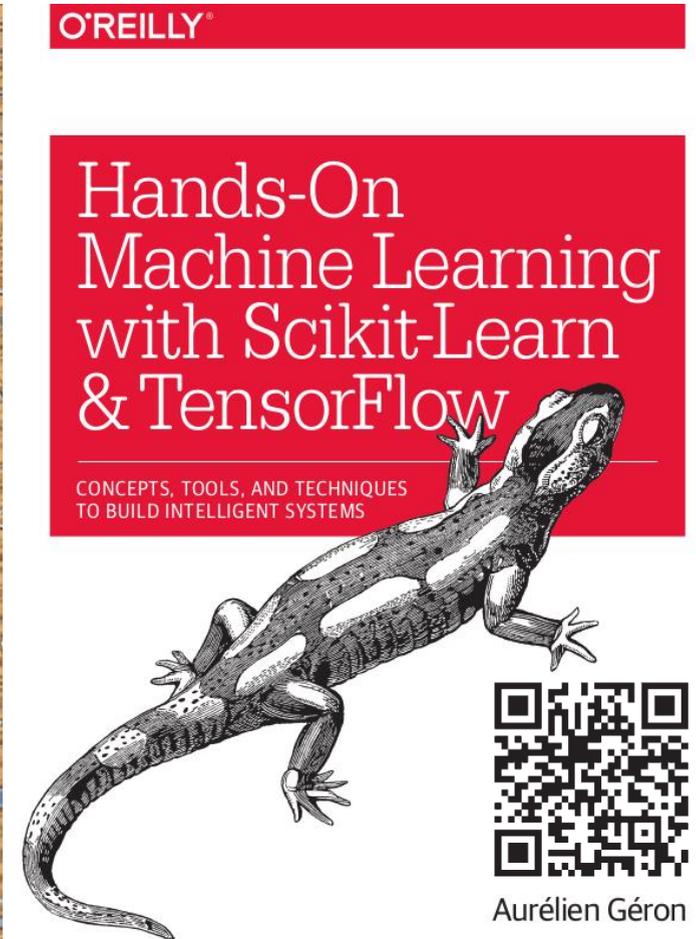
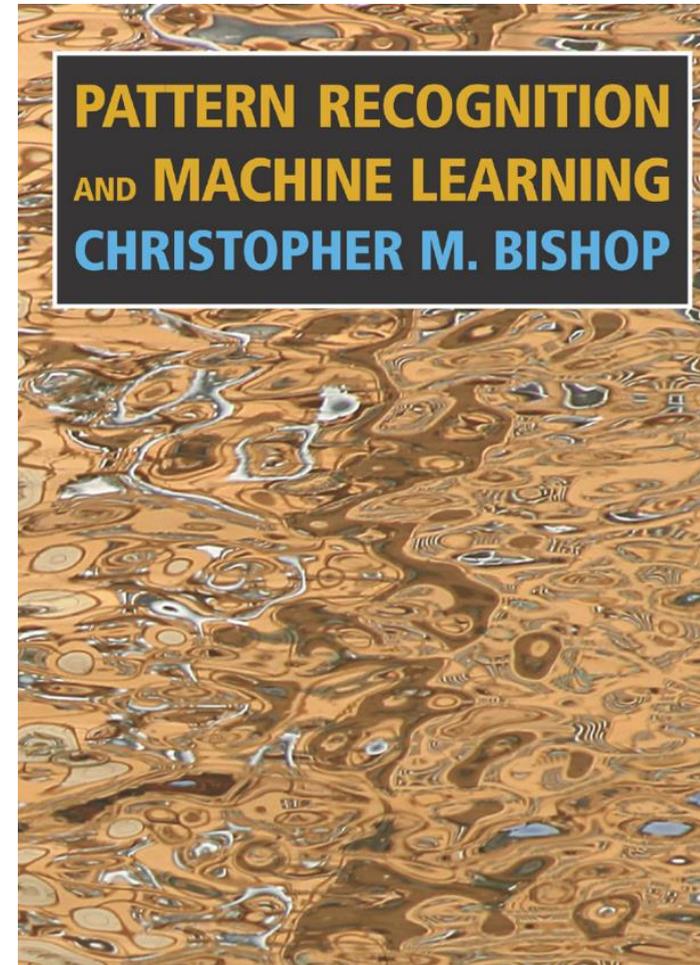
<https://github.com/Sampa-USP/scm2020/>

Introdução

Bishop C. M. Springer (2006), Geron A. O'Reilly (2017)

AM supervisionado

- Vetores de entrada e saída
- Classificadores
- Regressores
- Mínimos quadrados
- Gradiente descendente
- Redes neurais
- Aprendizado profundo
- (...)



Introdução

<http://courses.csail.mit.edu/6.036/> -- <http://courses.csail.mit.edu/6.862/>

- **1943: Walter Pitts and Warren McCulloch**
Computational models for neural nets
- **1957, 1962: Frank Rosenblatt**
Perceptron (Single-Layer & Multi-Layer)
- **1965: Alexey Ivakhnenko and V. G. Lapa**
Learning algorithm for MLP
- **1970: Seppo Linnainmaa**
Backpropagation and automatic differentiation
- **1979: Kunihiro Fukushima**
Convolutional neural networks
- **1982: John Hopfield**
Hopfield networks (recurrent neural networks)



AM em ciência de materiais

Propriedades modeladas por AM:

Table 1. Summary of material properties predicted with machine learning methods and corresponding references

Property	References
Curie temperature	31,283–287
Vibrational free energy and entropy	288
Band gap	40,41,132,159,283,289–300
Dielectric breakdown strength	38,44,45
Lattice parameter	300
Debye temperature and heat capacity	41–43
Glass transition temperature	301,302
Thermal expansion coefficient	41
Thermal boundary resistance	303
Thermal conductivity	37,46–51,304,305
Local magnetic moments	127,306
Melting temperature	39,48,307
Magnetocaloric effects	283
Grain boundaries	308

Grain boundary energy	309–312
Grain boundary mobility	312
Interface energy	300
Seebeck coefficient	46,313,314
Thermoelectric figure of merit	315
Bulk and shear moduli	40–42,132,184,185,316
Electrical resistivity	46
Density of states	109,317,318
Fermi energy and Poisson ratio	40
Dopant solution energy	319
Metal–insulator classification	65
Topological invariants	320–326
Superconducting critical temperature	73,76,122,327–329
Li-ion conductivity and battery state-of-charge	65,330,331

J. Schmidt et al., npj Computational Materials (2019) 83

Aprendizado & redes neurais

Origens do aprendizado de máquina supervisionado:

- McCulloch & Pitts (1943); Rosenblatt (1957).



Frank Rosenblatt
by Wikipedia

Psychological Review
Vol. 65, No. 6, 1958

THE PERCEPTRON: A PROBABILISTIC MODEL FOR
INFORMATION STORAGE AND ORGANIZATION
IN THE BRAIN¹

F. ROSENBLATT

Cornell Aeronautical Laboratory

Aprendizado & redes neurais

Aprendizado?



Frank Rosenblatt
by Wikipedia

- Como detectar a informação?
 - Estímulo e resposta; i.e. pares de dados, entradas e saídas correspondentes;
- Como armazenar e recuperar informações?
 - Entradas são enviadas a um neurón, são processadas e resultam em uma saída. i.e. neuróns possuem uma memória.

Rosenblatt, F. *Psyc. Rev.* 65 (6) 1957

Aprendizado & redes neurais

Aprendizado?



Frank Rosenblatt
by Wikipedia

Como informações armazenadas podem influenciar o no processamento/reconhecimento de novas informações?

- Feedback baseado no resultado (saída); neurons são sensíveis ao reforço positivo e negativo (back-propagation);
- Perceptrons podem aprender via tentativa e erro.

Rosenblatt, F. *Psyc. Rev.* 65 (6) 1957

Aprendizado & redes neurais

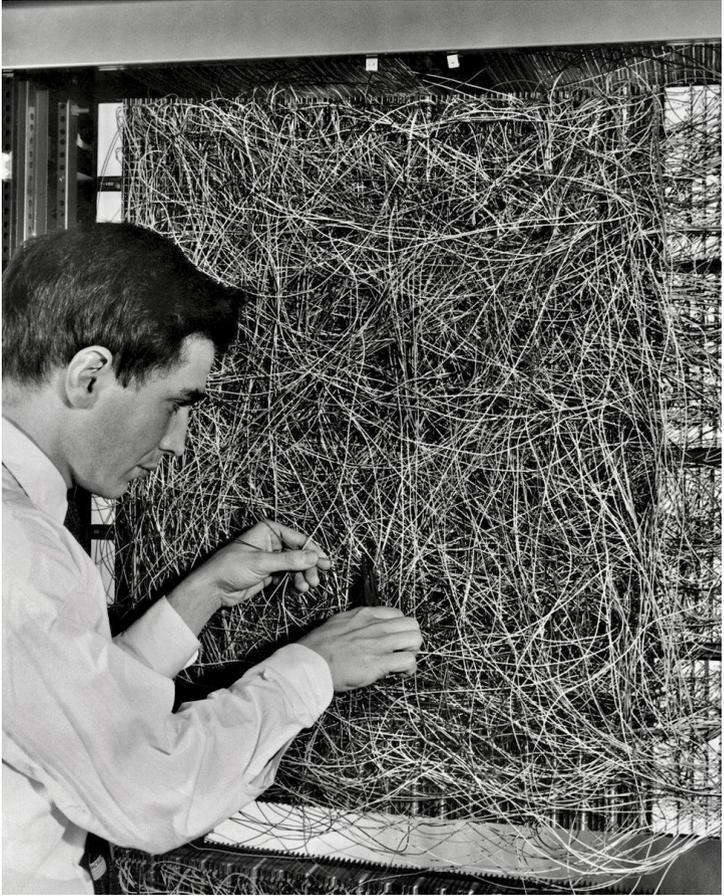
Aprendizado?



Frank Rosenblatt
by Wikipedia



[newyorker.com](https://www.newyorker.com)
[reddit.com](https://www.reddit.com)



Aprendizado de máquina

MIT course 6.036 - Lecture notes, Fall 2018

Aprendizado?



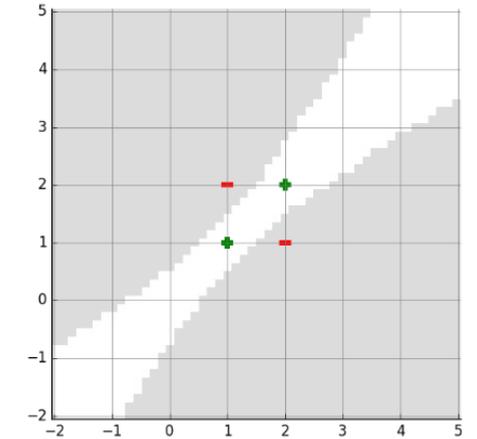
Frank Rosenblatt
by Wikipedia

Classificadores lineares

Teorema da convergência

Margem (erro de classificação)

Anos 1970: impossível modelar a função XOR



Anos 1980: aprendizado = problema de otimização;

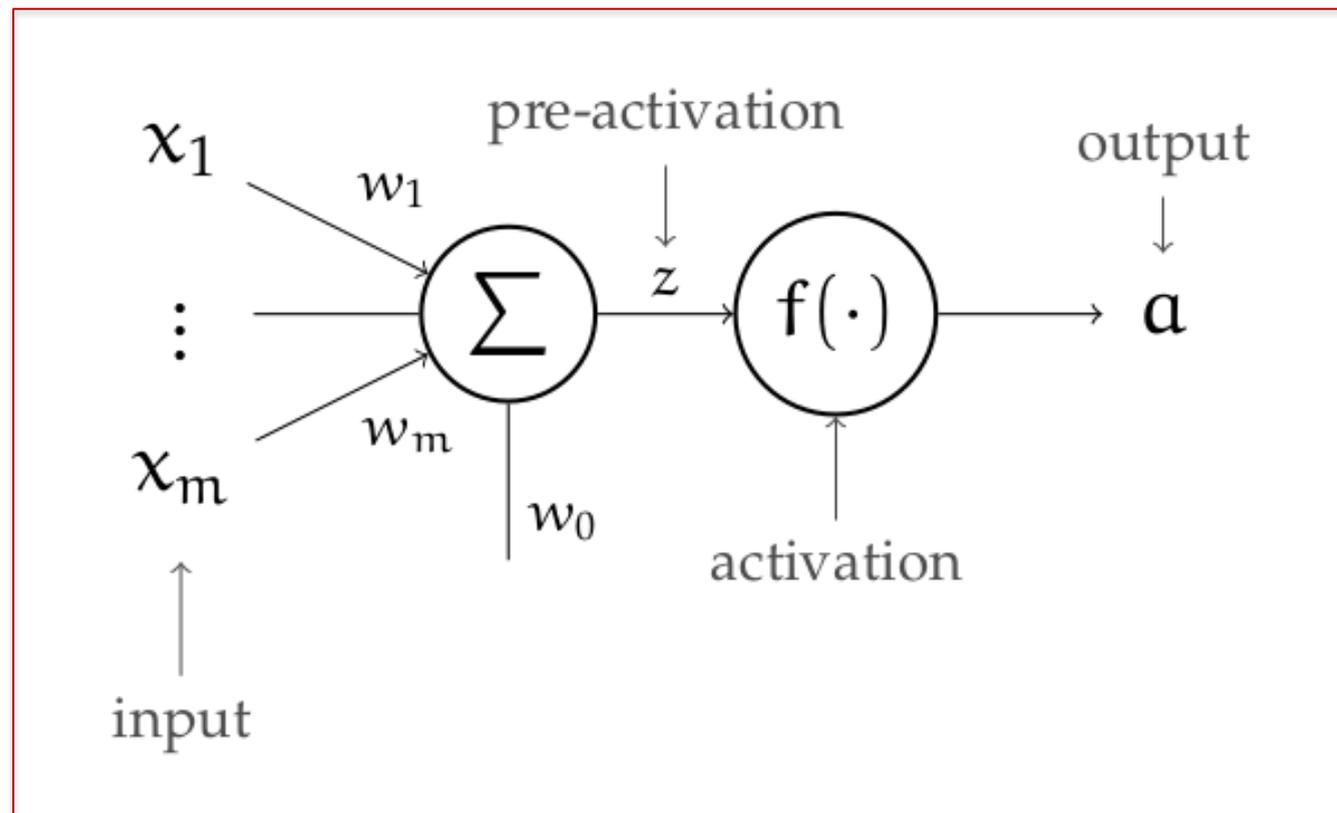
Gradiente descendente, etc.

$$J(\theta) = \left(\frac{1}{n} \sum_{i=1}^n \underbrace{L(x^{(i)}, y^{(i)}, \theta)}_{\text{loss}} \right)$$

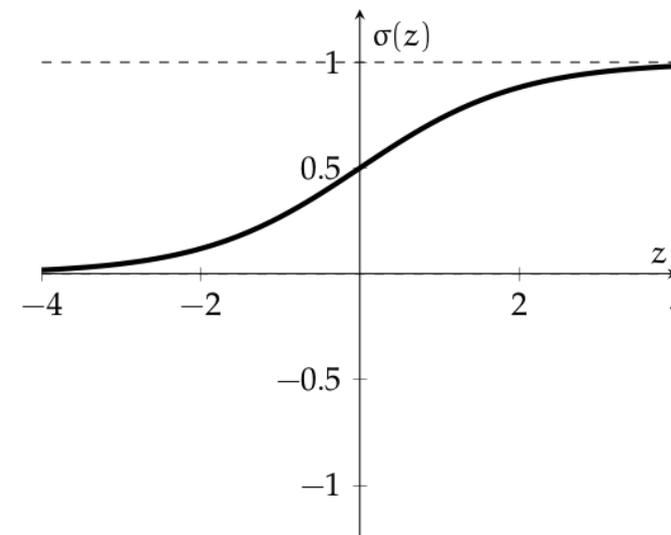
non-linearity
local min. traps
proc. limitations

Redes de um neurôn

- Funções de ativação - não linearidade (e.g. função logística)

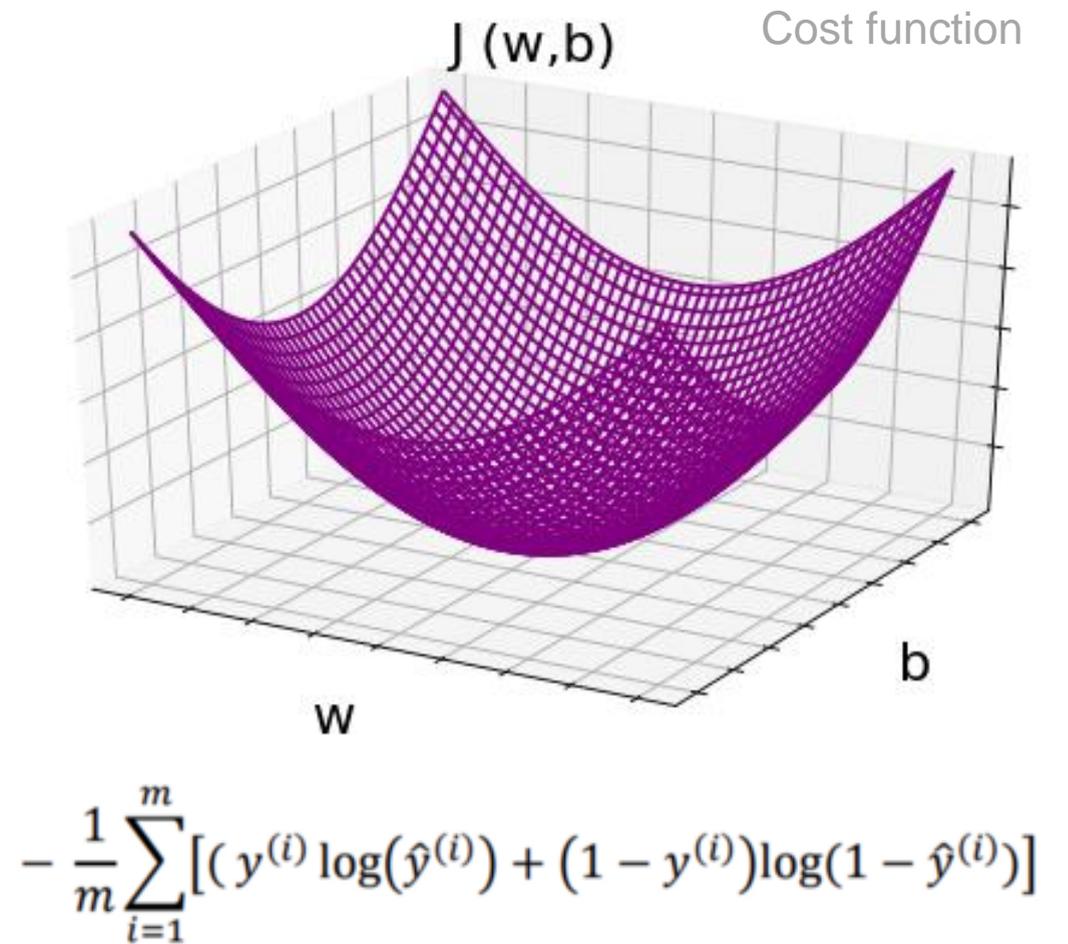
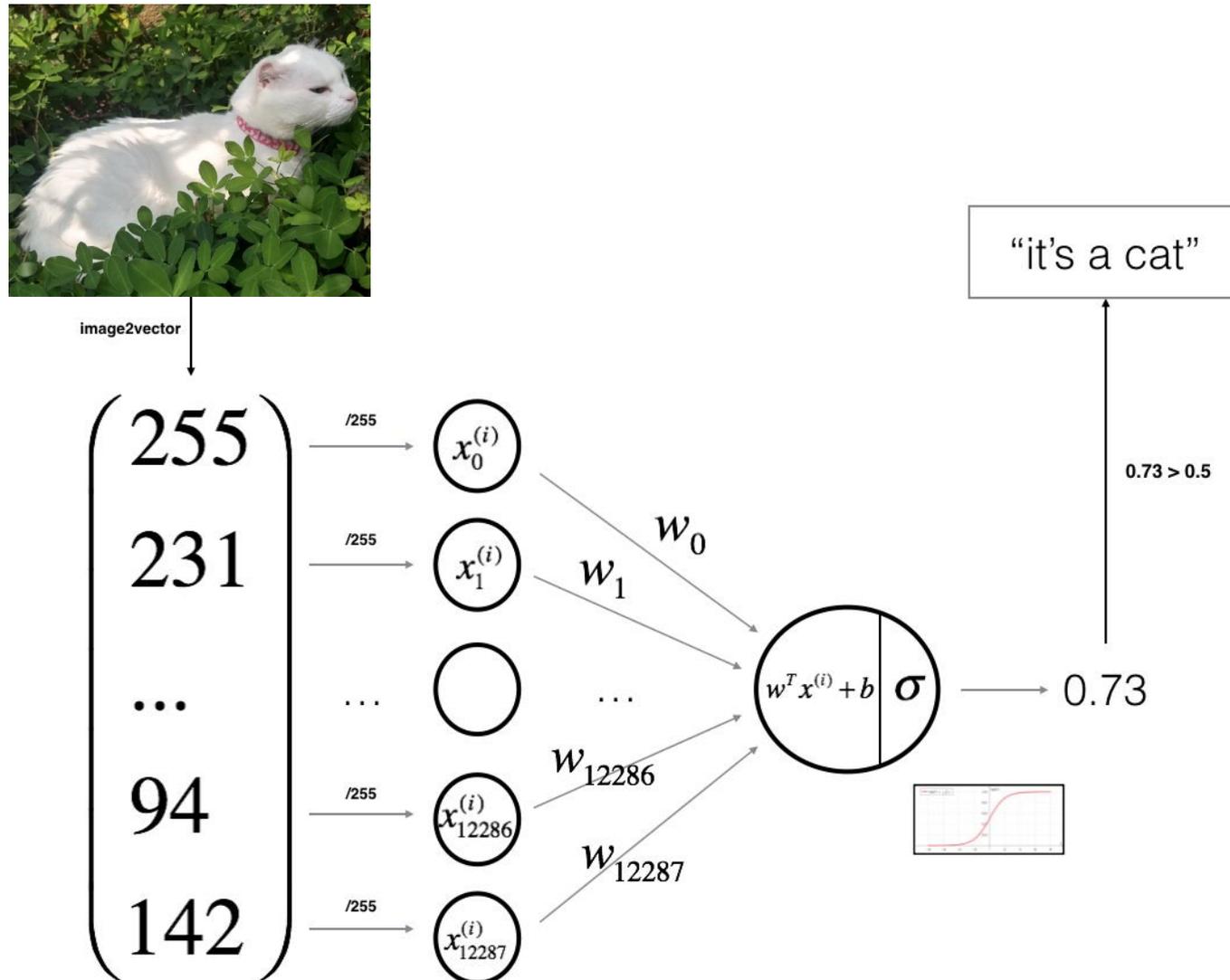


$$a = f(z) = f \left(\sum_{j=1}^m x_j w_j + w_0 \right) = f(\mathbf{w}^T \mathbf{x} + w_0) .$$



$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

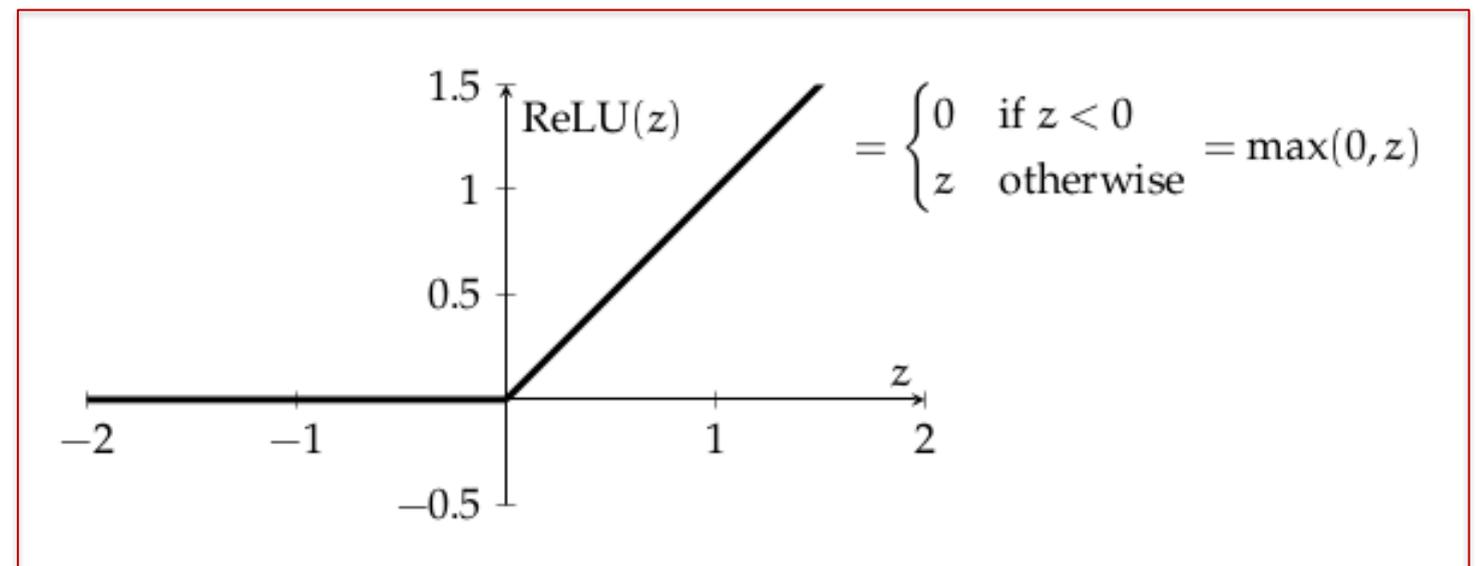
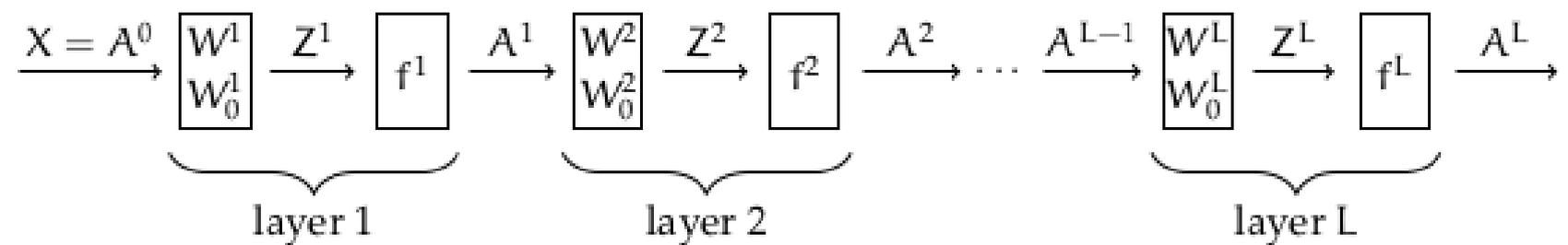
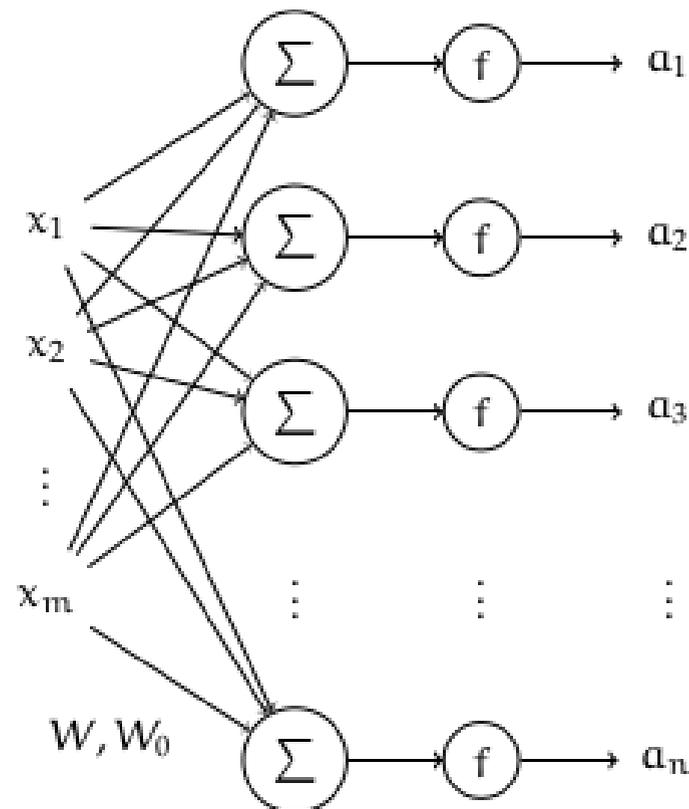
Redes neurais de uma camada



medium.com/@opetundeadepoju/
medium.com/@melodious

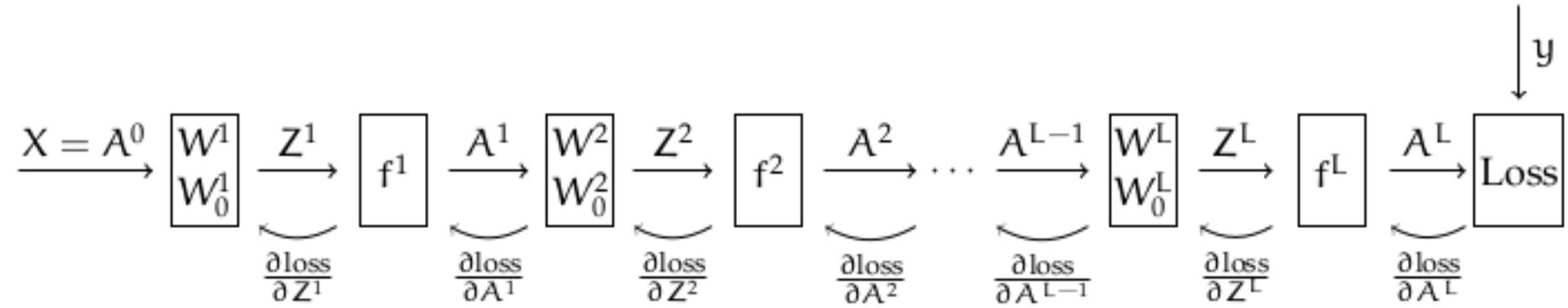
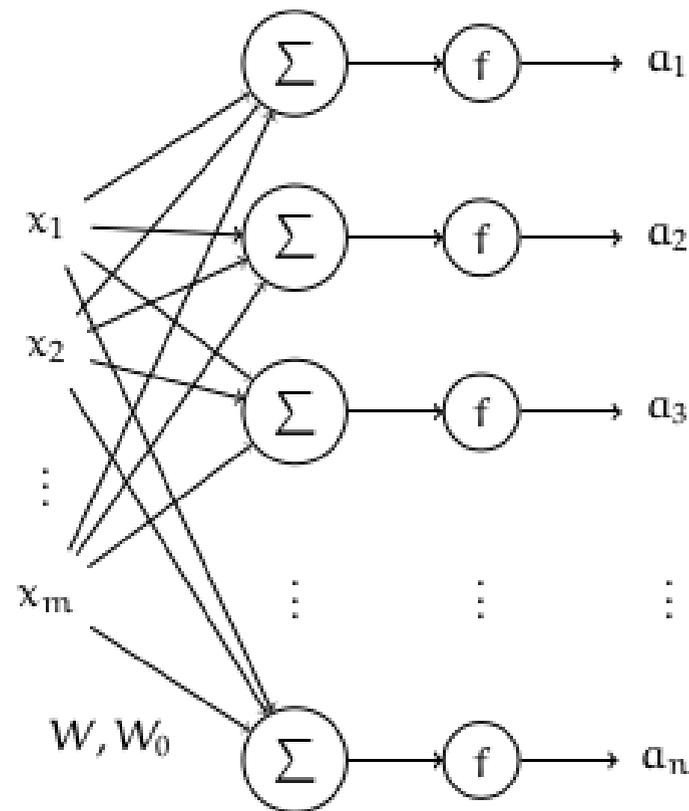
Redes sequenciais, múltiplas camadas

MIT course 6.036 - Lecture notes, Fall 2018

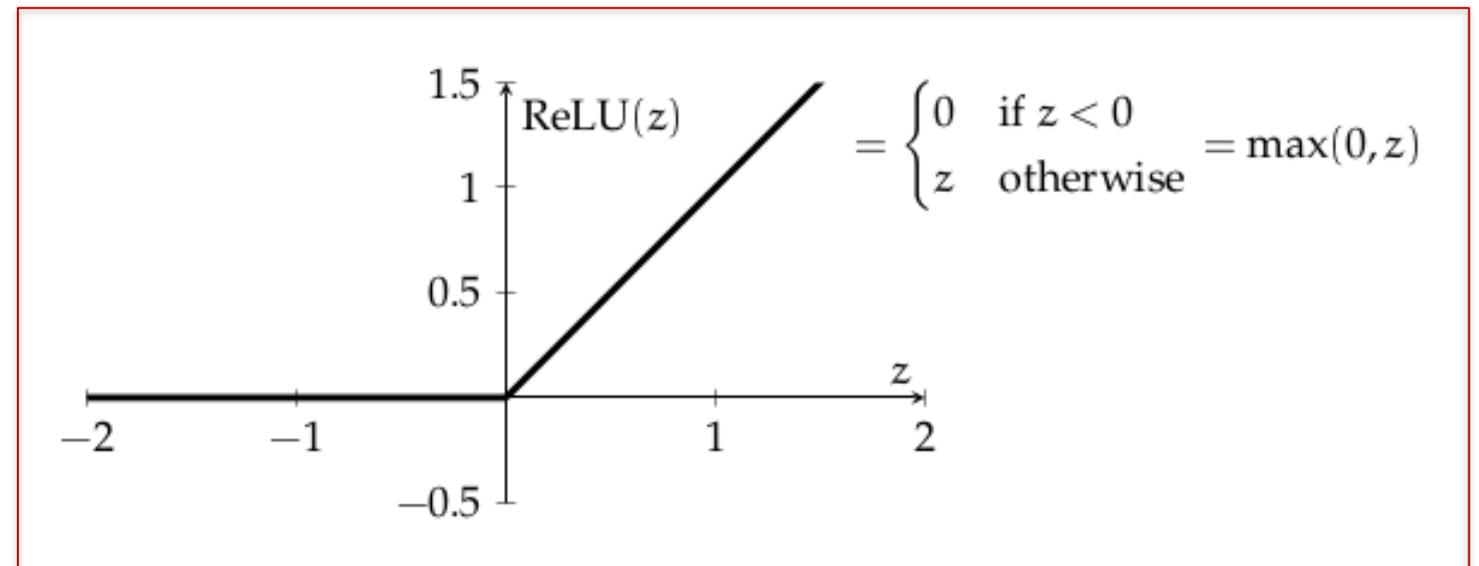


Redes profundas, L camadas

MIT course 6.036 - Lecture notes, Fall 2018

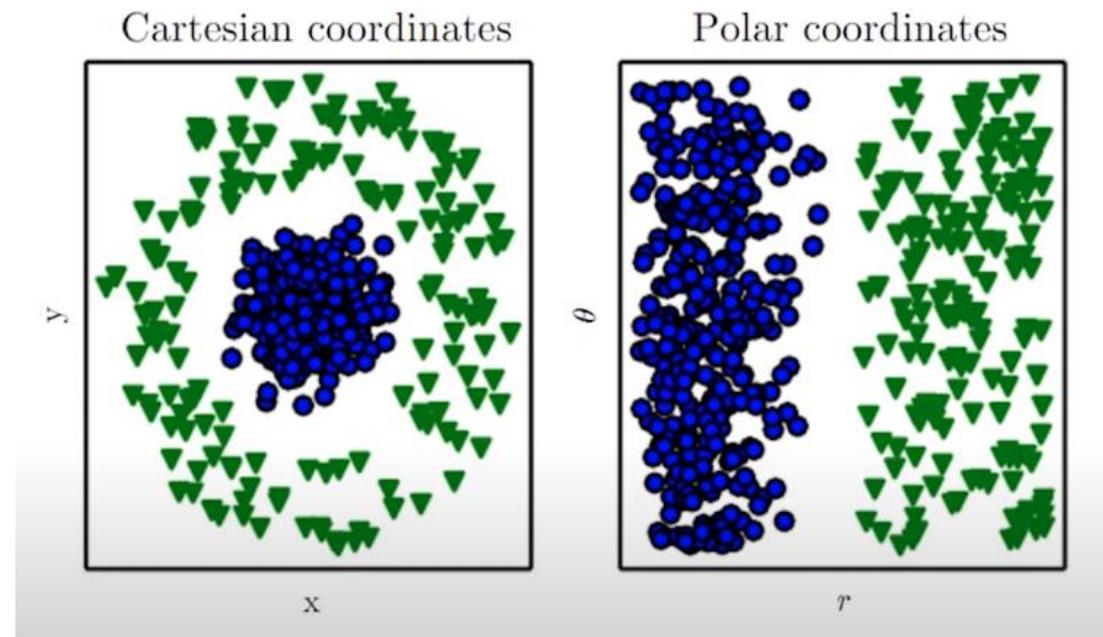


gradiente descendente [complexidade(L)]

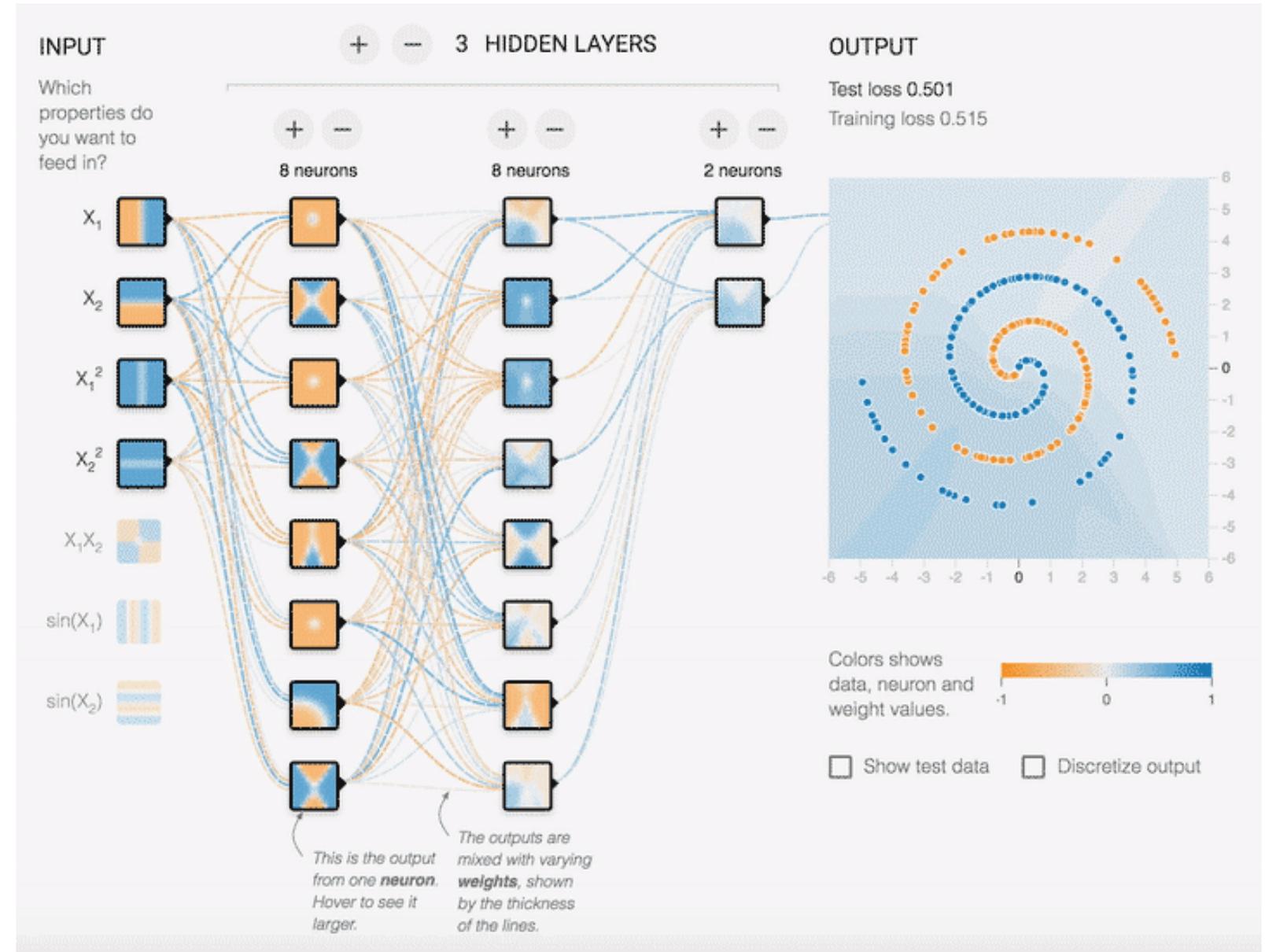


Redes profundas, L camadas

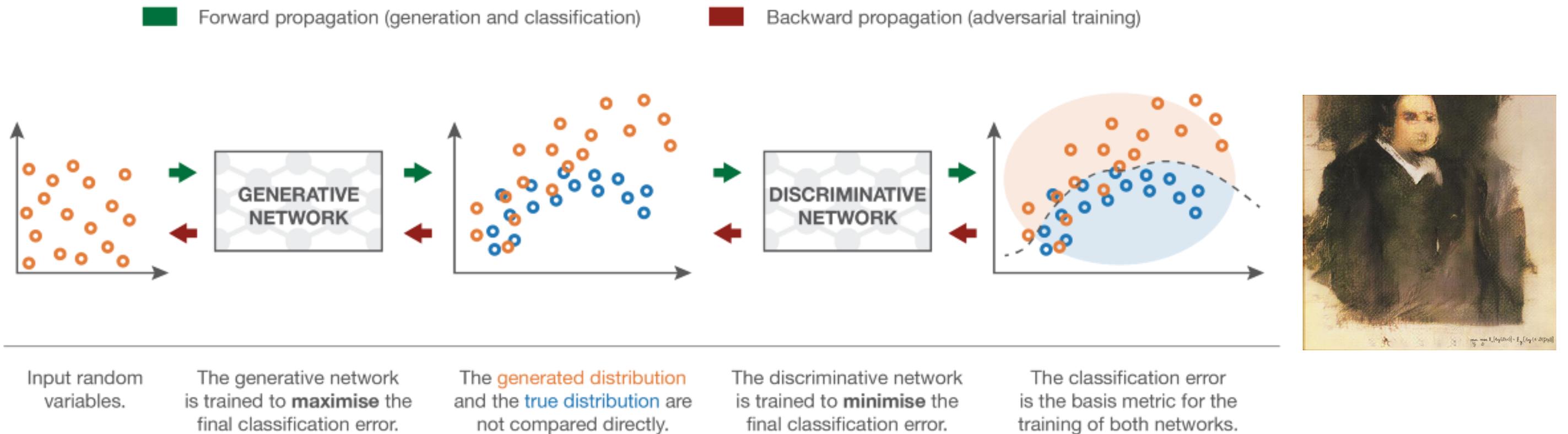
- Menos interferência humana



MIT - Deep Learning Basics

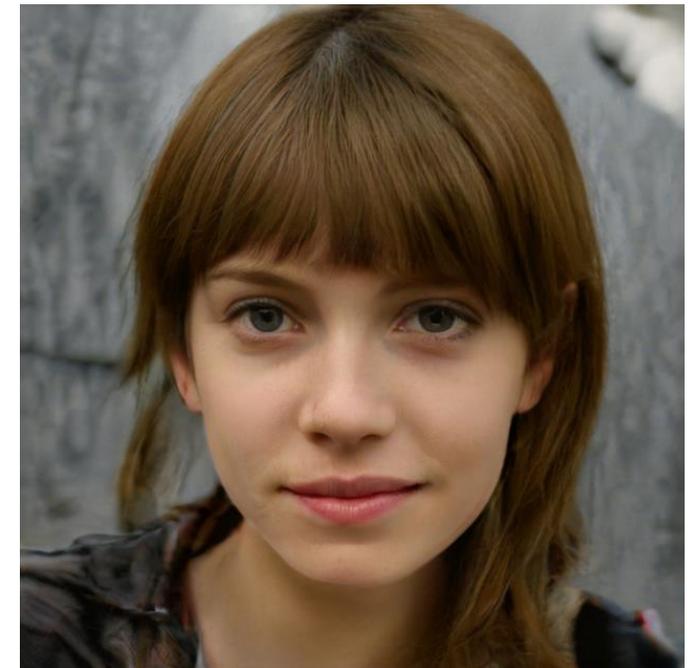
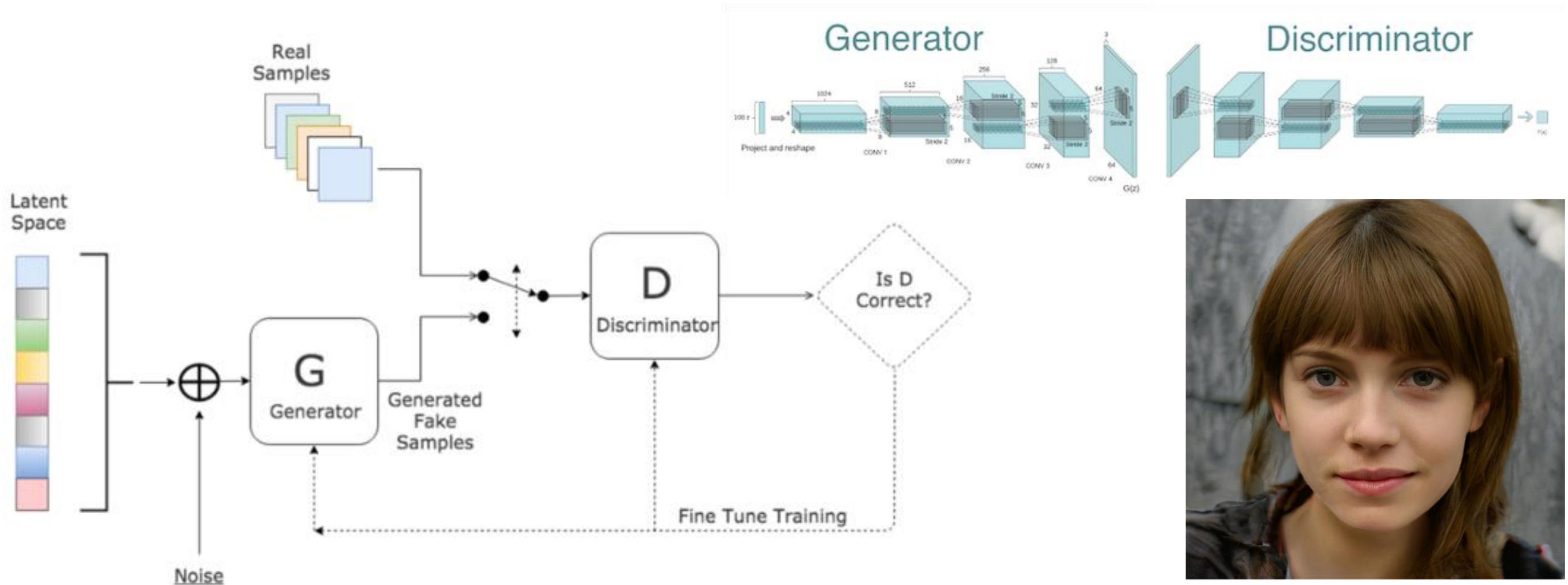


GANs (OpenAI era, 2014-)



<https://towardsdatascience.com/>

GANs (OpenAI era, 2014-)



<https://medium.com/machinelearningadvantage>
NVIDIA StyleGAN

SCM 2020

Prof. Caetano R. Miranda

IFUSP, São Paulo - Brazil

cmiranda@if.usp.br

Camilo A. F. Salvador

IFUSP, São Paulo - Brazil

csalvador@usp.br

APRENDIZADO DE MÁQUINA - TENDÊNCIAS

<https://github.com/Sampa-USP/scm2020/>

Redes neurais aprendem física (2020)

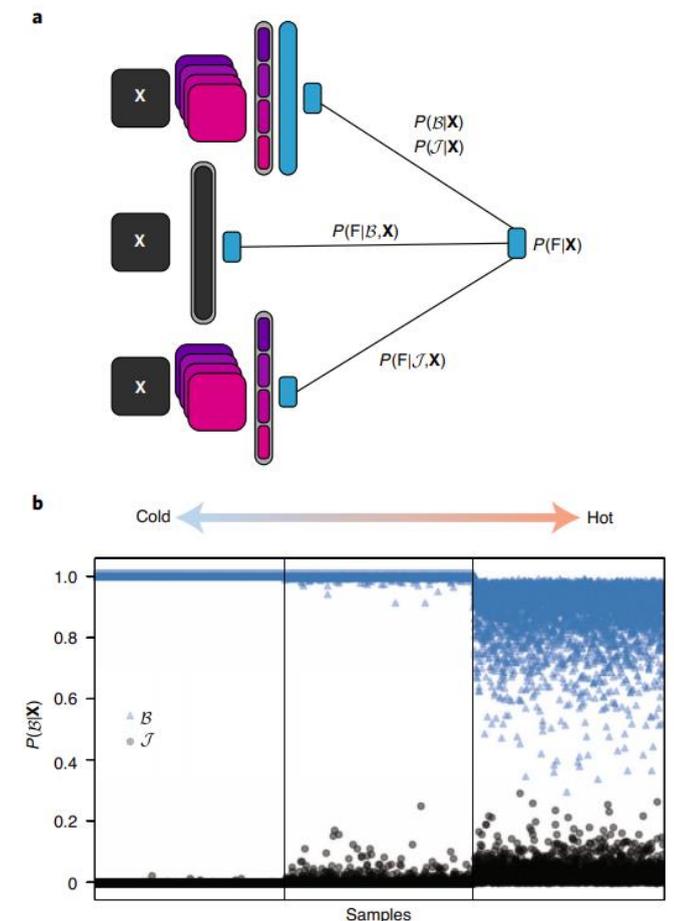


Check for updates

Machine learning the thermodynamic arrow of time

Alireza Seif ^{1,2} ✉, Mohammad Hafezi ^{1,2,3} and Christopher Jarzynski ^{1,4}

"...we have used various techniques to interpret what the network learns. In particular, by examining the optimized parameter values that emerge from the training, we have been able to identify which physical quantities the network uses to guess the direction of time's arrow. In this sense, **our study represents a step toward AI driven discovery of physical concepts.**"



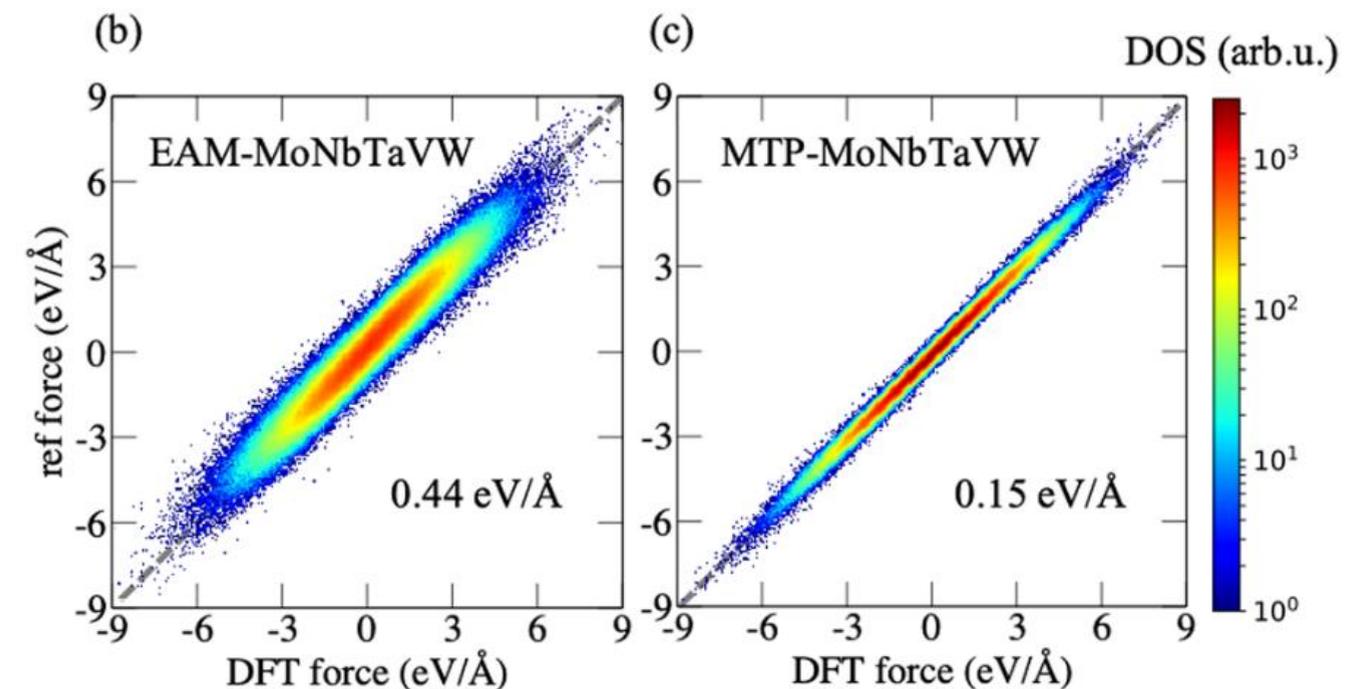
Seif et al., Nature Physics (in press) 2020

AM & Física de materiais

Potenciais interatômicos

Em função do grande número de **parâmetros ajustáveis**, potenciais baseados em AM são muito mais flexíveis, podendo **superar a acurácia** de potenciais clássicos ou campos de força.

- Figura: Correlação entre forças obtidas por DFT e potenciais (b) EAM (embedded-atom models) e (c) MTP (moment tensor potential), um método de aprendizado de máquina geométrico. Material: liga de alta entropia MoNbTaVW a 3000K

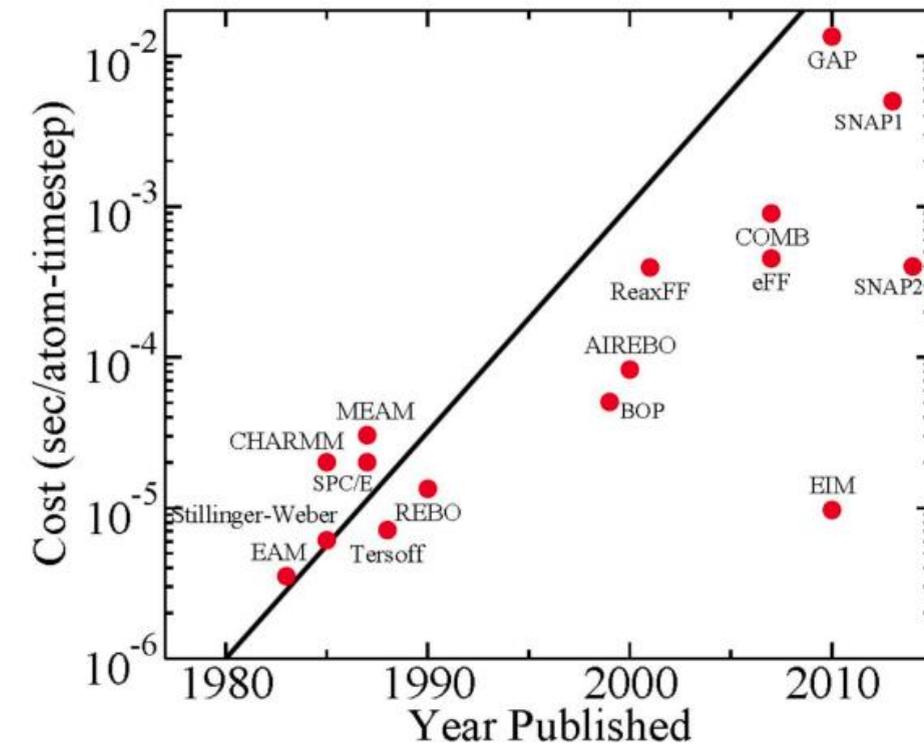


Reproduced from Grabowski et al., NPJ Comput. Mater. 5, 80 (2019). Copyright 2019 Author(s);

Ferrari et al., J. Appl. Phys. 128, 150901 (2020)

Por que obter novos potenciais?

Potential	System	Atoms	Timestep	CPU	LJ Ratio
Granular	chute flow	32000	0.0001 tau	5.08e-7	0.34x
FENE bead/spring	polymer melt	32000	0.012 tau	5.32e-7	0.36x
Lennard-Jones	LJ liquid	32000	0.005 tau	1.48e-6	1.0x
DPD	pure solvent	32000	0.04 tau	2.16e-6	1.46x
EAM	bulk Cu	32000	5 fmsec	3.59e-6	2.4x
Tersoff	bulk Si	32000	1 fmsec	6.01e-6	4.1x
Stillinger-Weber	bulk Si	32000	1 fmsec	6.10e-6	4.1x
EIM	crystalline NaCl	32000	0.5 fmsec	9.69e-6	6.5x
SPC/E	liquid water	36000	2 fmsec	1.43e-5	9.7x
CHARMM + PPPM	solvated protein	32000	2 fmsec	2.01e-5	13.6x
MEAM	bulk Ni	32000	5 fmsec	2.31e-5	15.6x
Peridynamics	glass fracture	32000	22.2 nsec	2.42e-5	16.4x
Gay-Berne	ellipsoid mixture	32768	0.002 tau	4.09e-5	28.3x
AIREBO	polyethylene	32640	0.5 fmsec	8.09e-5	54.7x
COMB	crystalline SiO2	32400	0.2 fmsec	4.19e-4	284x
eFF	H plasma	32000	0.001 fmsec	4.52e-4	306x
ReaxFF	PETN crystal	16240	0.1 fmsec	4.99e-4	337x
ReaxFF/C	PETN crystal	32480	0.1 fmsec	2.73e-4	185x
VASP/small	water	192/512	0.3 fmsec	26.2	17.7e6
VASP/medium	CO2	192/1024	0.8 fmsec	252	170e6
VASP/large	Xe	432/3456	2.0 fmsec	1344	908e6



Moore's Law for Interatomic Potentials
Plimpton and Thompson, MRS Bulletin (2012).

lammps.sandia.gov/

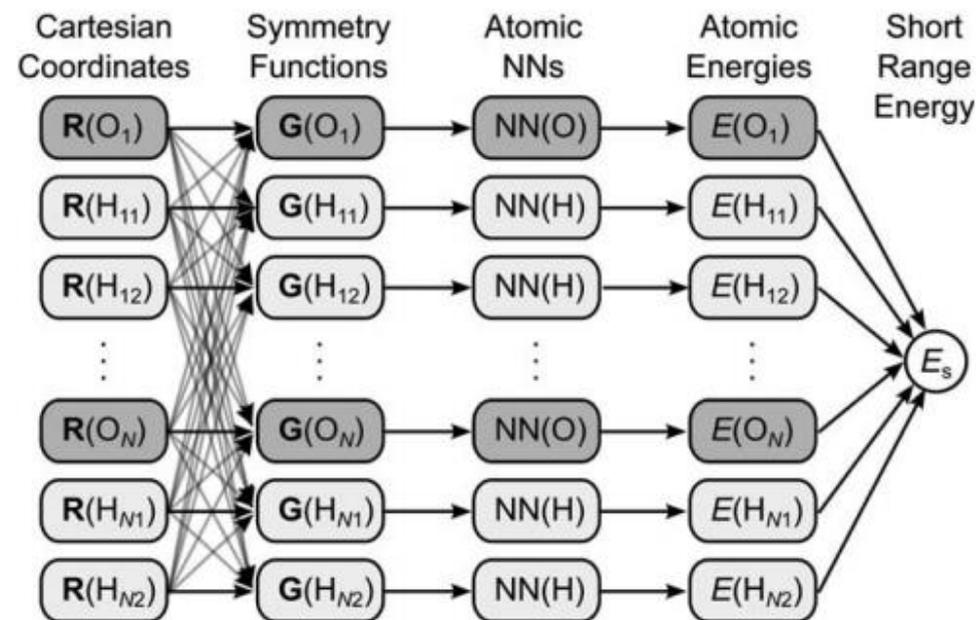
Behler & Parrinello

Funções de simetria (2007):

Funções que representam a vizinhança

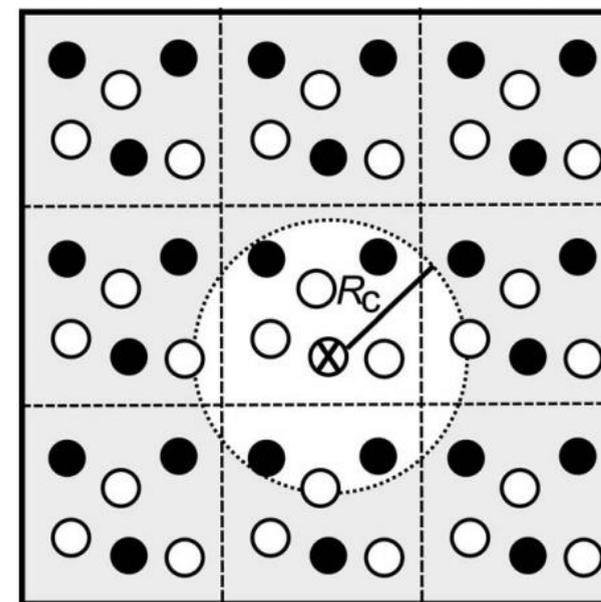
Invariantes a rotação e translação

Possibilitam obter potenciais transferíveis



$$f_{c,1}(R_{ij}) = \begin{cases} 0.5 \cdot \left[\cos\left(\frac{\pi R_{ij}}{R_c}\right) + 1 \right] & \text{for } R_{ij} \leq R_c \\ 0.0 & \text{for } R_{ij} > R_c \end{cases}$$

$$f_{c,2}(R_{ij}) = \begin{cases} \tanh^3 \left[1 - \frac{R_{ij}}{R_c} \right] & \text{for } R_{ij} \leq R_c \\ 0.0 & \text{for } R_{ij} > R_c \end{cases}$$



Behler, J., Int. Journal of Quantum Chemistry 2015, 115, 1032–1050

Potenciais baseados em redes neurais

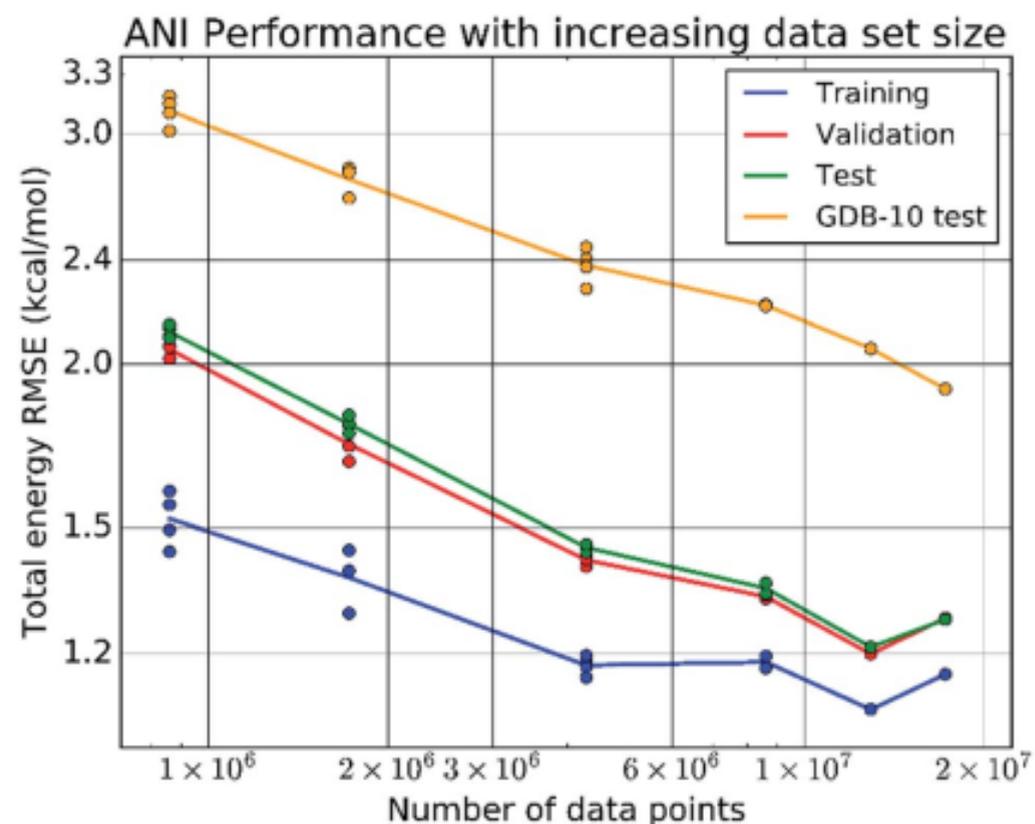


Fig. 3 Log-log plots of the training, validation, testing, and a random GDB-10 (molecules with 10 heavy atoms from the GDB-11 database) extensibility testing set of total energy errors vs. increasing number of data points in the training set. The sets of points converge to the final ANI-1 potential presented in this paper, trained on the full ANI-1 data set.

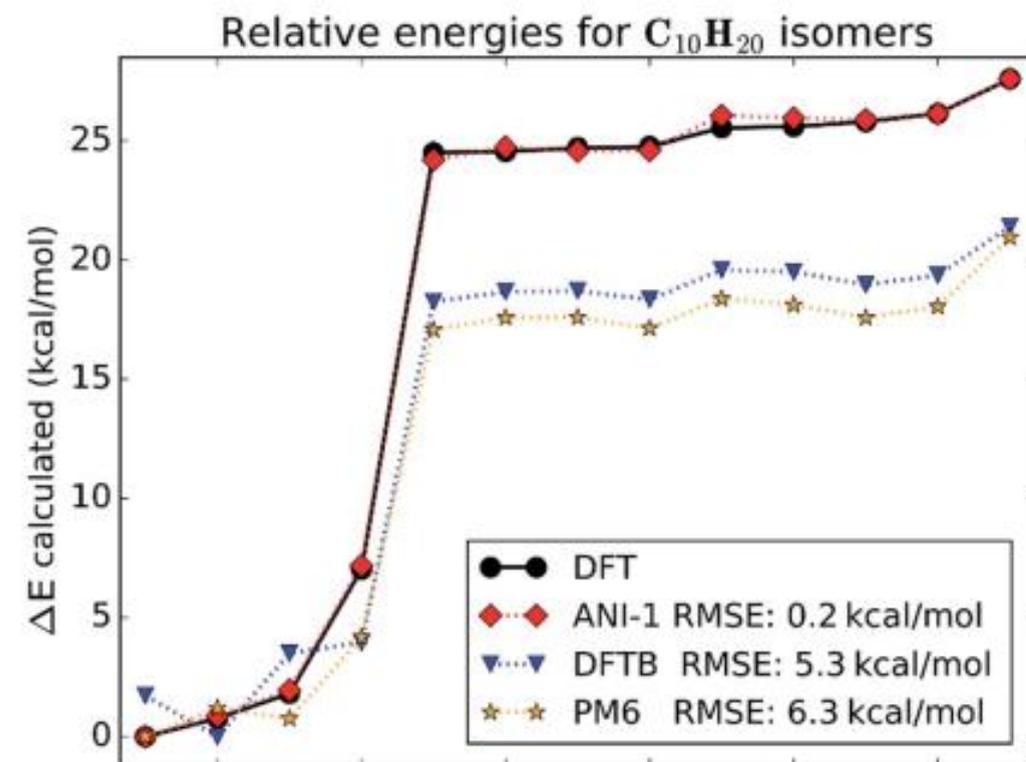
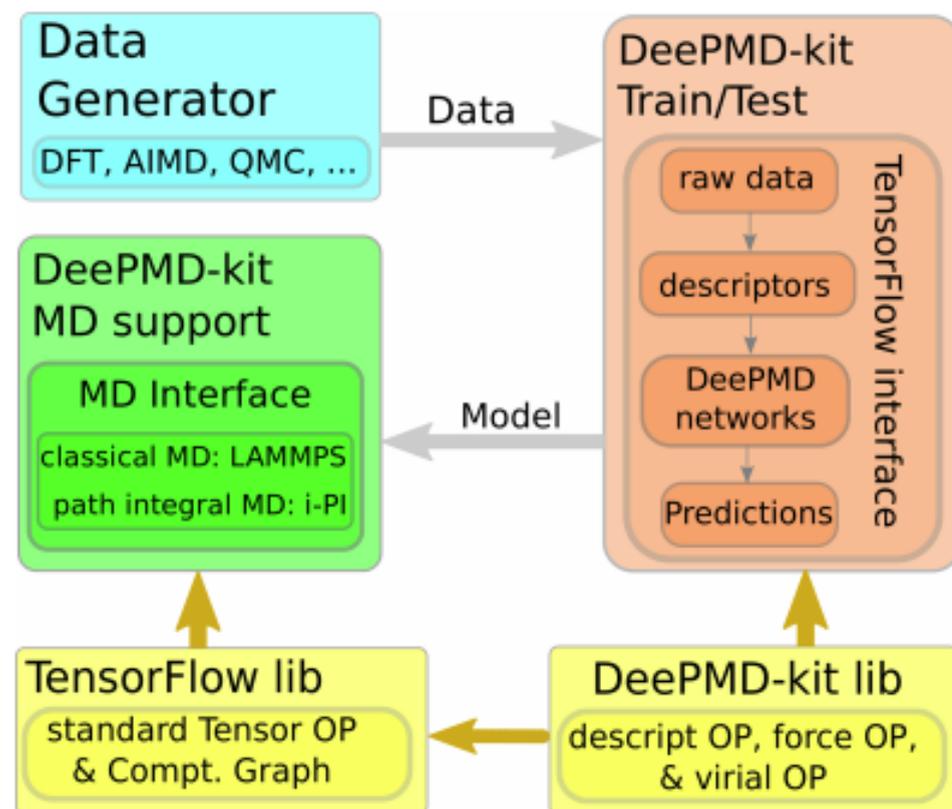


Fig. 5 The total energies, shifted such that the lowest is zero, calculated for various $C_{10}H_{20}$ isomers, are compared between DFT with the $\omega B97X$ functional and 6-31G(d) basis set, the ANI-1 potential, AM1 semi-empirical, and PM6 semi-empirical methods.

Smith et al., Chem. Sci., 2017, 8, 3192–3203

Potenciais interatômicos H-O

Pacote DeepMD



Variáveis descritivas são construídas com base nos átomos vizinhos (dentro de um raio arbitrário) preservando:

- Simetria de translação
- Simetria de rotação
- Simetria de permutação

$$D_i = D_i(\mathbf{R}_i, \{\mathbf{R}_j \mid j \in N_{R_c}(i)\}).$$

É uma transformação linear com base nas coordenadas atômicas

Função de ativação: tangente hiperbólica.

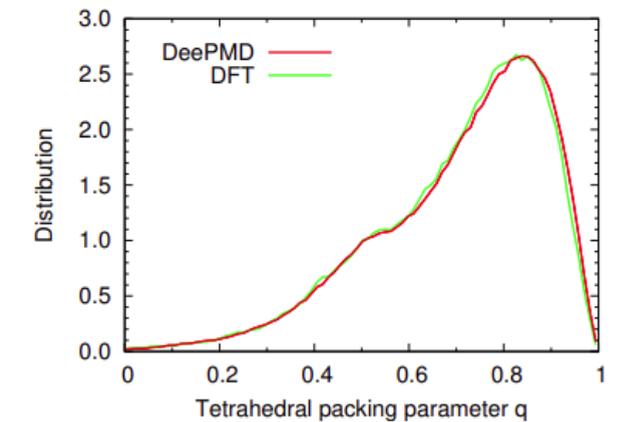
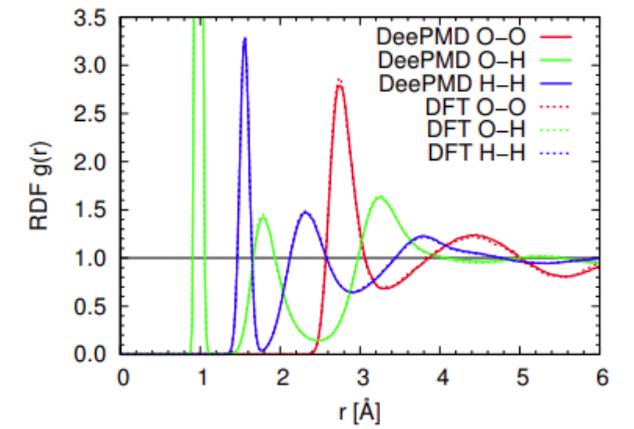
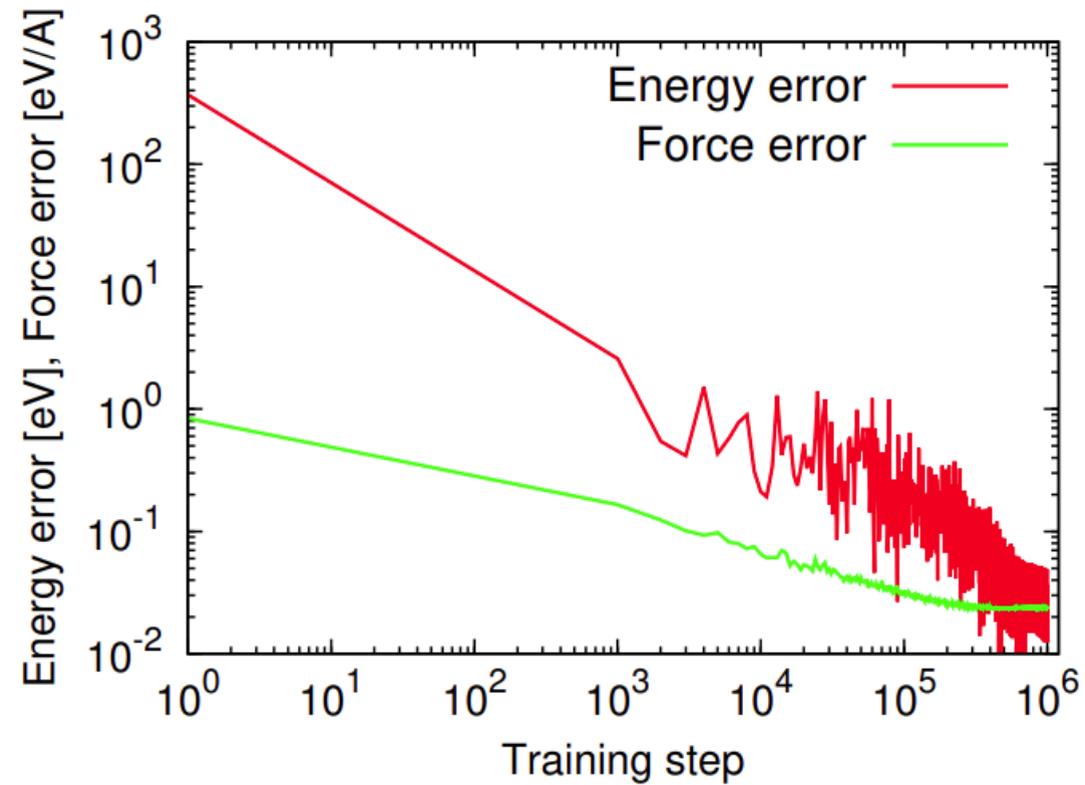
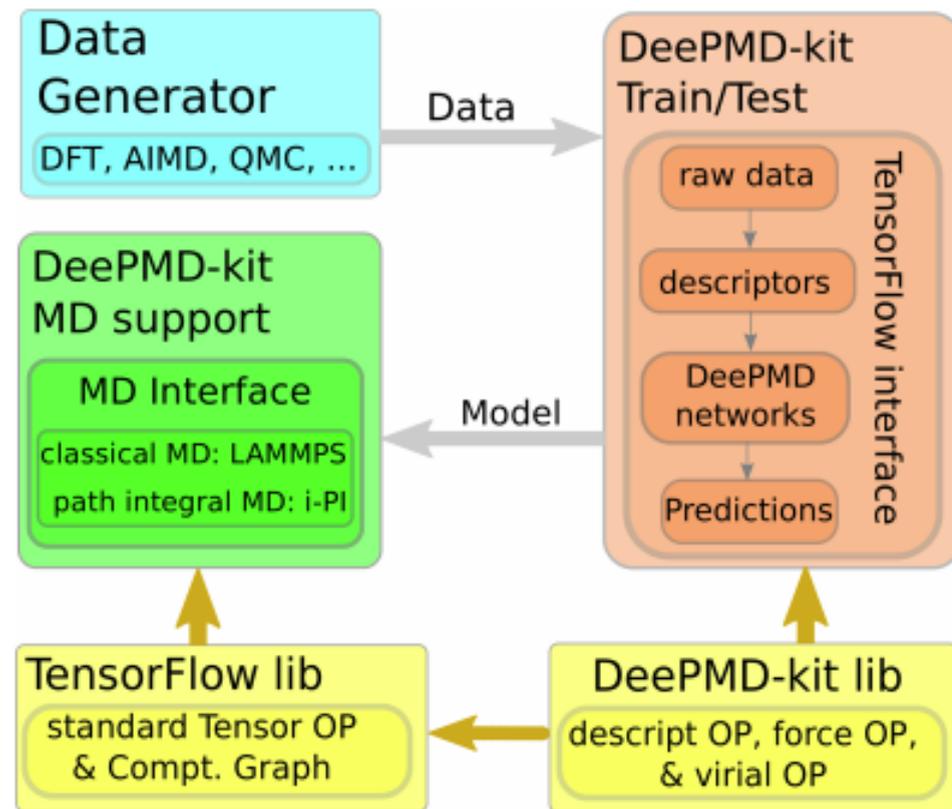
A função de perda é balanceada para forças e energias:

$$L(p_\epsilon, p_f, p_\xi) = \frac{p_\epsilon}{N} \Delta E^2 + \frac{p_f}{3N} \sum_i |\Delta \mathbf{F}_i|^2 + \frac{p_\xi}{9N} \|\Delta \Xi\|^2,$$

Wang et al., Comp. Phys. Sim. (228) 2018

Potenciais interatômicos H-O

Pacote DeepMD



RMSE 0.028 eV
RMSF 0.024(eV/Å)

Wang et al., Comp. Phys. Sim. (228) 2018

Potencial interatômico combinando DFT + ML

