

# APRENDIZADO NÃO SUPERVISIONADO

## APRESENTAÇÃO DA DISCIPLINA



# Quem sou eu?



Thomas Paula

✉️ [tsp.thomas@gmail.com](mailto:tsp.thomas@gmail.com)  
🐦 [@tsp\\_thomas](https://twitter.com/tsp_thomas)



PUCRS

Pontifícia Universidade Católica  
do Rio Grande do Sul

Machine Learning  
PORTO ALEGRE



# E vocês, quem são? :)

---

- ✗ Qual seu nome?
- ✗ Com o que você trabalha/o que estuda?
- ✗ Tem experiência com Machine Learning?
- ✗ Por que procuraram a especialização? Qual o objetivo?
- ✗ Quais suas expectativas para a disciplina?



# Tópicos da Disciplina

---

- ✗ Introdução ao Aprendizado Não Supervisionado
- ✗ Agrupamento de Dados (Clustering)
- ✗ Redução de Dimensionalidade
- ✗ Modelos Generativos



# Trabalhos

---

- ✗ Vamos ter algumas implementações a serem entregues
- ✗ Será dado tempo em aula para trabalhar e tirar dúvidas
- ✗ Todos os aspectos serão considerados
  - ✗ Código e documentação
  - ✗ Experimentos realizados e conclusões
  - ✗ Entrega na data combinada
- ✗ Entrega via Github (repositório acessível publicamente) ou email



# Calendário (sujeito a alterações)

09/01/19	Introdução ao Aprendizado Não Supervisionado
10/01/19	Agrupamento de Dados
15/01/19	Agrupamento de Dados
16/01/19	Agrupamento de Dados
17/01/19	Redução de Dimensionalidade
22/01/19	Modelos Generativos e Fechamento da Disciplina



# Machine Learning



what society thinks I  
do

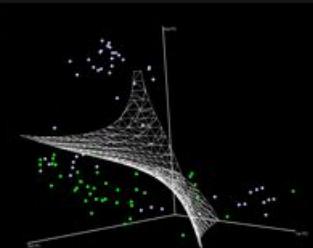


what my friends think  
I do



what my parents think  
I do

$$\begin{aligned}L_t &= \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^n \alpha_i y_i (\mathbf{x}_i \cdot \mathbf{w} + b) + \sum_{i=1}^n \alpha_i \\&\alpha_i \geq 0, \forall i \\&\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i, \sum_{i=1}^n \alpha_i y_i = 0 \\&\nabla \hat{g}(\theta_t) = \frac{1}{n} \sum_{i=1}^n \nabla \ell(x_i, y_i; \theta_t) + \nabla r(\theta_t). \\&\theta_{t+1} = \theta_t - \eta_t \nabla \ell(x_{i(t)}, y_{i(t)}; \theta_t) - \eta_t \cdot \nabla r(\theta_t) \\&\mathbb{E}_{i(t)}[\ell(x_{i(t)}, y_{i(t)}; \theta_t)] = \frac{1}{n} \sum_i \ell(x_i, y_i; \theta_t).\end{aligned}$$



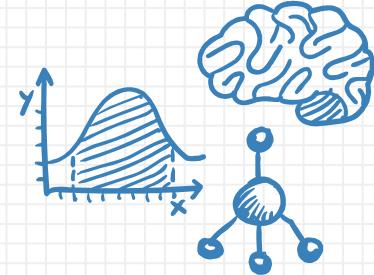
```
>>> from sklearn import svm
```

what other programmers  
think I do

what I think I do

what I really do





# INTRODUÇÃO AO APRENDIZADO NÃO SUPERVISIONADO



# Agenda de Hoje

---

- ✗ Introdução
  - ✗ Recapitulando Aprendizado Supervisionado
  - ✗ Exemplos do estado-da-arte
- ✗ Aprendizado Não Supervisionado
  - ✗ Definições e exemplos
  - ✗ Exemplos do estado-da-arte



# Introdução

Recapitulando Aprendizado Supervisionado



# Aprendizado Supervisionado

---

- ✗ O que é Aprendizado Supervisionado?
- ✗ Quais algoritmos são exemplos de Aprendizado Supervisionado?
- ✗ Como podemos avaliar algoritmos de Aprendizado Supervisionado?
- ✗ O que é overfitting?
- ✗ Qual é o principal objetivo de Machine Learning?

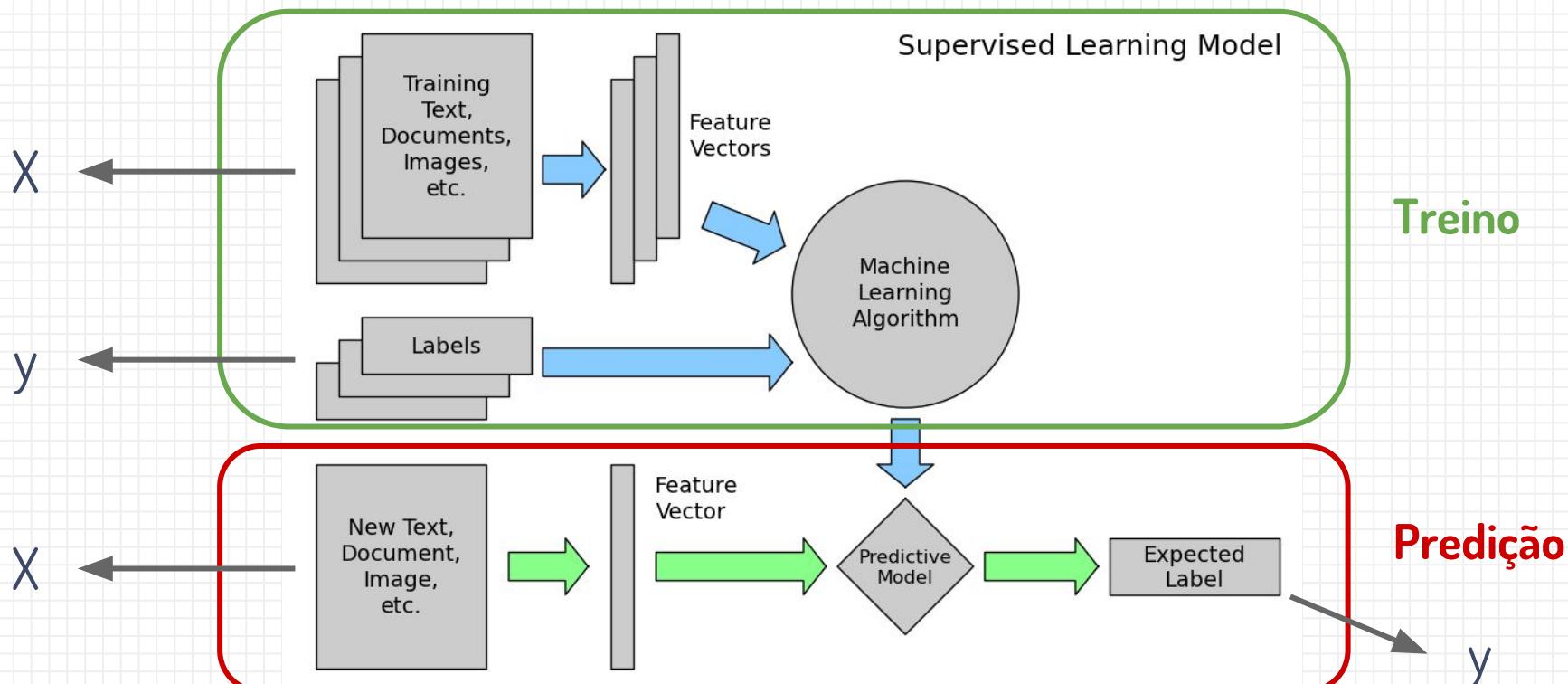


# Aprendizado Supervisionado

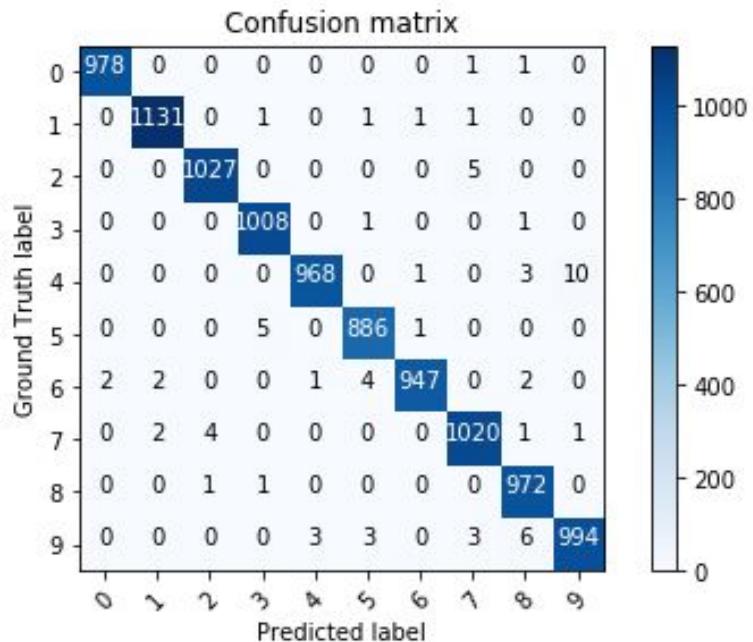
- ✗ Aprender mapeamento de uma entrada para uma saída
- ✗ Tem um “professor externo”
  - ✗ Respostas são fornecidas (dados rotulados)



# Aprendizado Supervisionado



# Como avaliar?



# Aprendizado Supervisionado

Alguns exemplos do estado-da-arte



# Semantic Segmentation

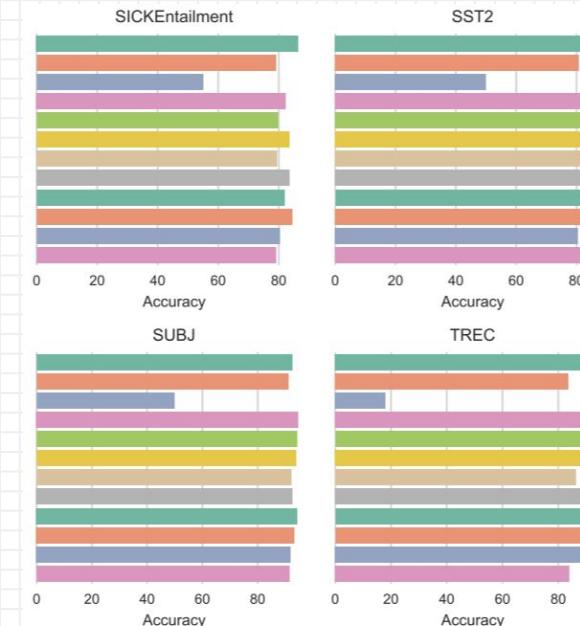


Fonte: He, Kaiming, et al. "Mask r-cnn." Computer Vision (ICCV), 2017 IEEE International Conference on. IEEE, 2017.



# Natural Language Processing (NLP)

Dataset	Task	Example	Output
Customer Reviews (CR) [19]	Sentiment analysis of customer products' reviews	We tried it out Christmas night and it worked great .	Positive
Multi-Perspective Question and Answering (MPQA) [39]	Evaluation of opinion polarity	Don't want	Negative
Movie Reviews (MR) [31]	Sentiment analysis of movie reviews	Too slow for a younger crowd , too shallow for an older one .	Negative
Stanford Sentiment Analysis 2 (SST-2) [36]	Sentiment analysis with two classes: Negative and Positive	Audrey Tautou has a knack for picking roles that magnify her [...]	Positive
Stanford Sentiment Analysis 5 (SST-5) [36]	Sentiment analysis with 5 classes, that range from 0 (most negative) to 5 (most positive)	Nothing about this movie works	0
Subjectivity / Objectivity (SUBJ) [30]	Classify the sentence as Subjective or Objective	A movie that doesn't aim too high , but doesn't need to .	Subjective
Text REtrieval Conference (TREC) [38]	Question and answering	What are the twin cities ?	LOC:city



Fonte: Christian S. Perone, Roberto Silveira, and Thomas S. Paula. "Evaluation of sentence embeddings in downstream and linguistic probing tasks." arXiv preprint arXiv:1806.06259(2018).



# Image Captioning



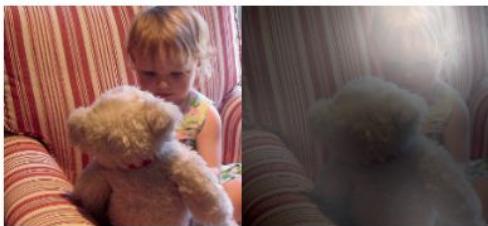
A woman is throwing a **frisbee** in a park.



A **dog** is standing on a hardwood floor.



A **stop** sign is on a road with a mountain in the background



A little **girl** sitting on a bed with a teddy bear.



A group of **people** sitting on a boat in the water.



A giraffe standing in a forest with **trees** in the background.

Fonte: LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521.7553 (2015): 436.



# Speech Recognition

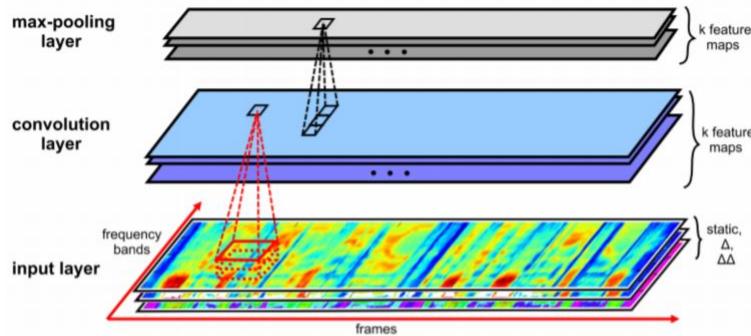


Figure 1: The convolution layer and max-pooling layer applied upon input features.

Model	NP	Dev PER	Test PER
BiLSTM-3L-250H [12]	3.8M	-	18.6%
BiLSTM-5L-250H [12]	6.8M	-	18.4%
TRANS-3L-250H [12]	4.3M	-	18.3%
CNN-(3,5)-10L-ReLU	4.3M	17.4%	19.3%
CNN-(3,5)-10L-PReLU	4.3M	17.2%	18.9%
CNN-(3,5)-6L-maxout	4.3M	18.7%	21.2%
CNN-(3,5)-8L-maxout	4.3M	17.7%	19.8%
CNN-(3,3)-10L-maxout	4.3M	18.4%	19.9%
CNN-(3,5)-10L-maxout	4.3M	<b>16.7%</b>	<b>18.2%</b>

Fonte: Zhang, Ying, et al. "Towards end-to-end speech recognition with deep convolutional neural networks." arXiv preprint arXiv:1701.02720 (2017).



# Text Recognition



(a) ICDAR 2015



(b) ICDAR 2017 MLT

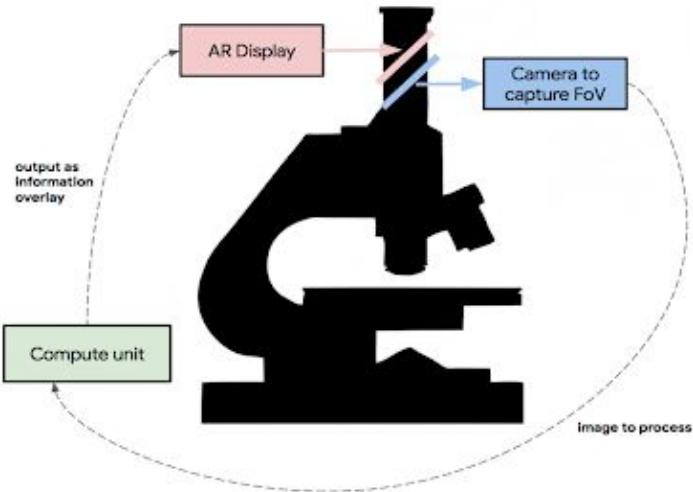


(c) ICDAR 2013

Fonte: Liu, Xuebo, et al. "FOTS: Fast Oriented Text Spotting with a Unified Network." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.



# Cancer Detection

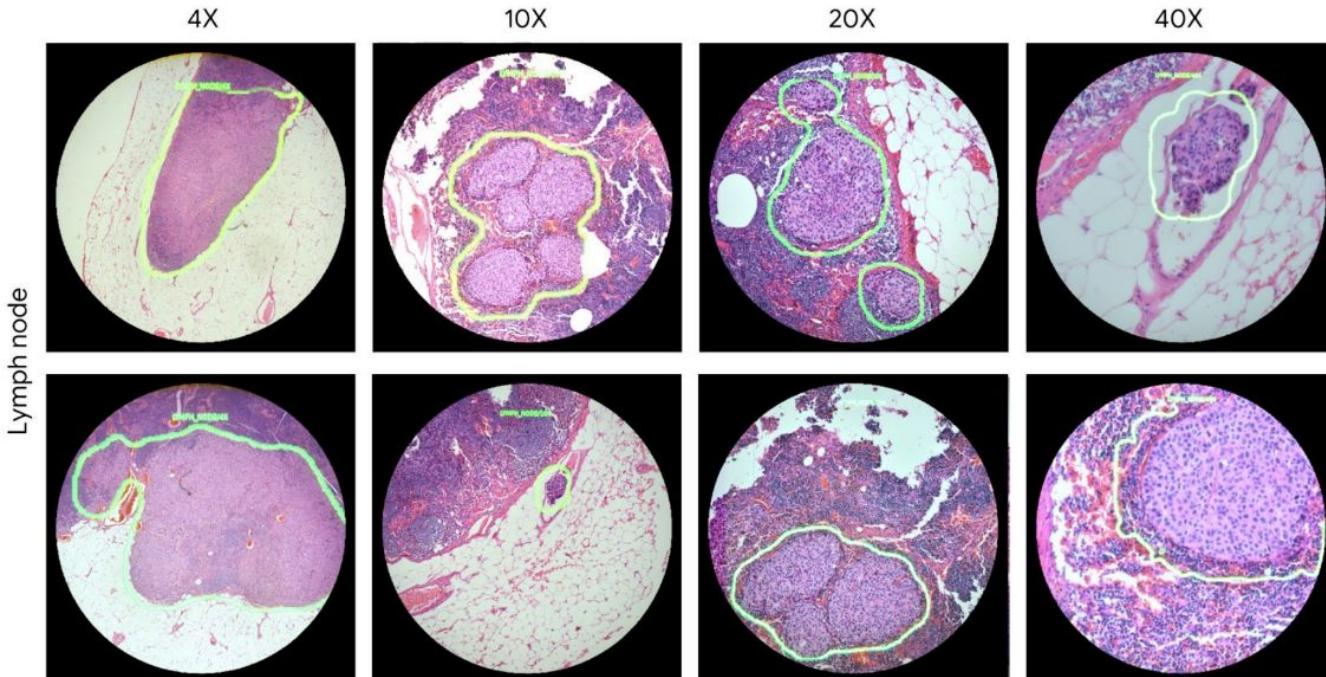


Research at Google

Fonte: An Augmented Reality Microscope, Google Research Blog, 2018.



# Cancer Detection



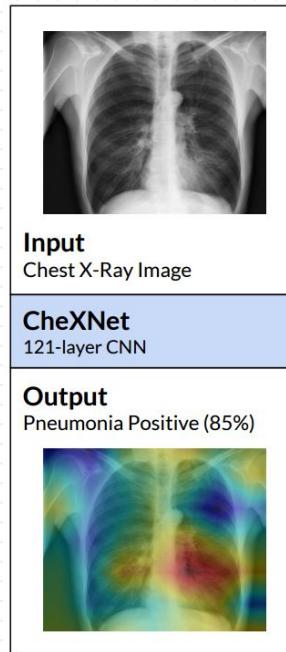
Fonte: An Augmented Reality Microscope, Google Research Blog, 2018.



# Pneumonia Detection on Chest X-Rays

## Pathology

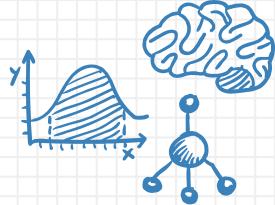
Atelectasis  
Cardiomegaly  
Effusion  
Infiltration  
Mass  
Nodule  
Pneumonia  
Pneumothorax  
Consolidation  
Edema  
Emphysema  
Fibrosis  
Pleural Thickening  
Hernia



	F1 Score (95% CI)
Radiologist 1	0.383 (0.309, 0.453)
Radiologist 2	0.356 (0.282, 0.428)
Radiologist 3	0.365 (0.291, 0.435)
Radiologist 4	0.442 (0.390, 0.492)
Radiologist Avg.	0.387 (0.330, 0.442)
CheXNet	0.435 (0.387, 0.481)

Fonte: Rajpurkar, Pranav, et al. "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning." (2017).





Tá, mas com resultados tão bacanas...  
por que queremos Aprendizado Não  
Supervisionado?



# O que há de errado?

---

- ✗ Necessita de dados rotulados
  - ✗ No caso de Deep Learning, precisa de muitos mesmo!
  - ✗ Rotular dados é geralmente um processo caro e que consome tempo
  
- ✗ Em geral, bastante conectado com a tarefa onde foi treinado
  - ✗ Aprendizado deve ser generalizável para outros domínios
  - ✗ Temos problemas ao apresentar instâncias muito diferentes das vistas durante o treinamento



# O que há de errado? (cont'd)

---

## Unsupervised Deep Learning - Alex Graves (NeurIPS 2018)

- ✗ Se nosso objetivo é criar sistemas inteligentes que tenham sucesso em uma grande variedade de tarefas, porque não ensinamos tais tarefas diretamente?
  
- ✗ Alvos (rótulos) podem ser difíceis de obter
- ✗ Aprendizado Não Supervisionado “parece mais humano”
- ✗ Queremos generalização rápida para novas tarefas e situações

Fonte: Unsupervised Deep Learning Tutorial, Alex Graves and Marc'Aurelio Ranzato, NeurIPS 2018.



# Machine Learning Cake



**Yann Lecun**

Chief Scientist of  
Facebook AI Research

- **"Pure" Reinforcement Learning (cherry)**
  - ▶ The machine predicts a scalar reward given once in a while.
  - ▶ **A few bits for some samples**
  
- **Supervised Learning (icing)**
  - ▶ The machine predicts a category or a few numbers for each input
  - ▶ Predicting human-supplied data
  - ▶ **10→10,000 bits per sample**
  
- **Unsupervised/Predictive Learning (cake)**
  - ▶ The machine predicts any part of its input for any observed part.
  - ▶ Predicts future frames in videos
  - ▶ **Millions of bits per sample**



# Aprendizado Não Supervisionado



# Aprendizado Não Supervisionado

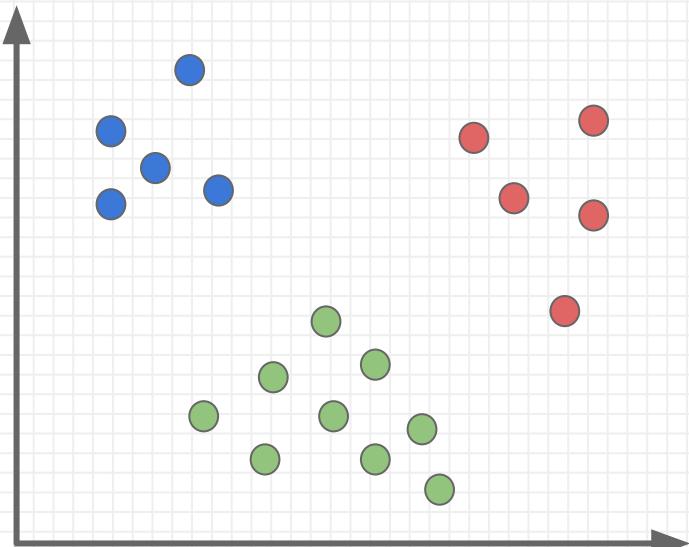
---

- ✗ Paradigma de Machine Learning que aprende a partir de dados não rotulados;
- ✗ Busca encontrar regularidades/padrões nos dados
  - ✗ Existe uma estrutura no espaço dos dados que certos padrões ocorrem mais frequentemente do que outros
  - ✗ Quer se entender porque determinados padrões acontecem mais do que outros

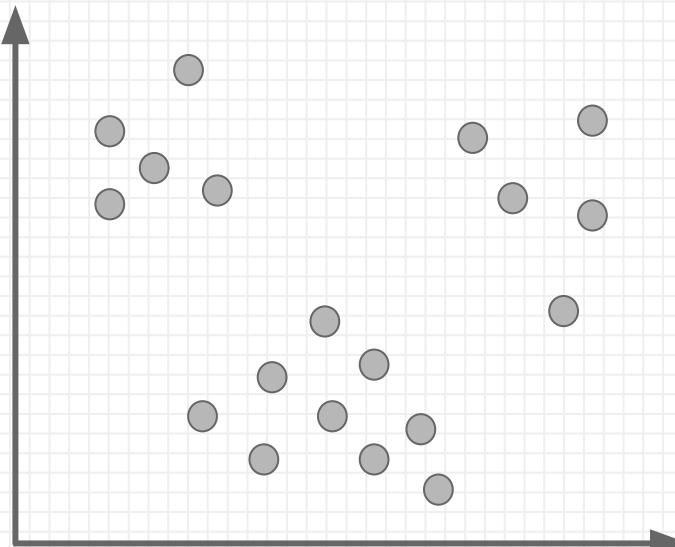
Fonte: Alpaydin, Ethem. Introduction to machine learning. MIT press, 2009.



# Aprendizado Não Supervisionado



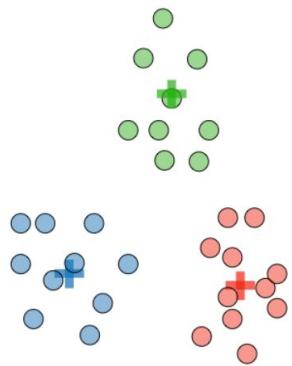
Supervisionado



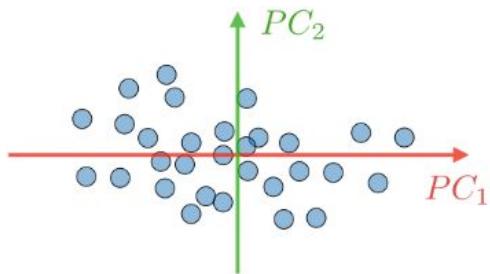
Não Supervisionado



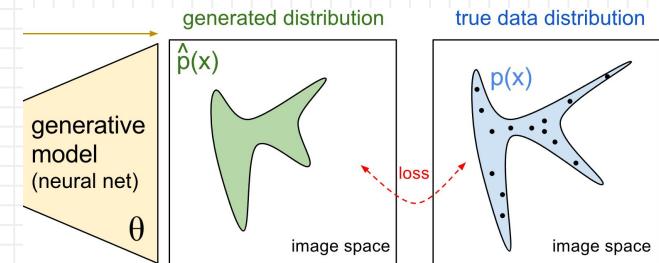
# Aprendizado Não Supervisionado



Agrupamento de Dados  
(Clustering)



Redução de  
Dimensionalidade



Modelos Generativos



# Diferenciando de outros paradigmas

## Semi-supervised

- ✗ Parte dos dados são rotulados e parte dos dados não são
- ✗ Ao se utilizar pequena parte dos dados rotulados, pode-se obter resultados melhores do que apenas não supervisionado
- ✗ Além disso, “custo” reduzido quando comparado com Aprendizado Supervisionado



# Diferenciando de outros paradigmas

## Semi-supervised

- ✗ Exemplo: Pseudo-Labeling
  - 1. Treinamento com dados rotulados
  - 2. Classificação dos dados não rotulados
  - 3. Treinamento com todos dados



# Diferenciando de outros paradigmas

## Weakly supervised

- ✗ Incomplete supervision: semi-supervised pode ser considerado um caso;
- ✗ Inexact supervision: dados rotulados de forma “superficial” (*coarse*)
- ✗ Inaccurate supervision: *ground truth* não totalmente correto (e.g. rótulos incorretos)

Fonte: Zhou, Zhi-Hua. "A brief introduction to weakly supervised learning." National Science Review 5.1 (2017): 44–53.

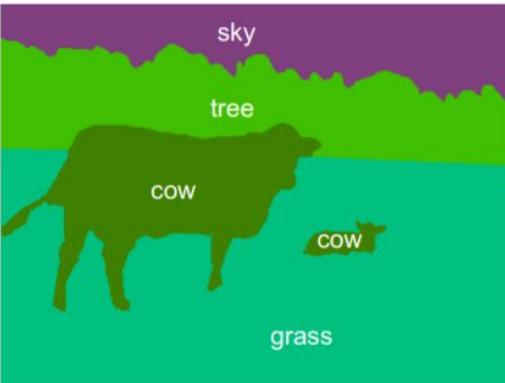


# Diferenciando de outros paradigmas

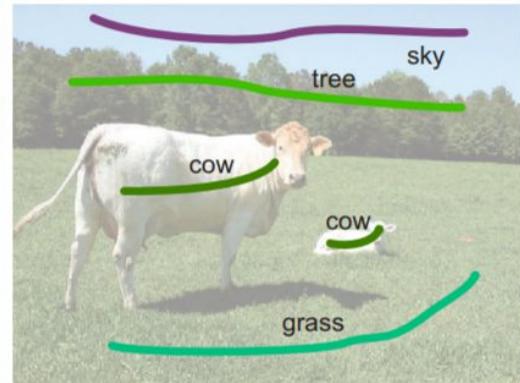
## Weakly supervised



(a) image



(b) mask annotation



(c) scribble annotation

Fonte: Lin, Di, et al. "Scribblesup: Scribble-supervised convolutional networks for semantic segmentation." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.



# Diferenciando de outros paradigmas

## Self-supervised

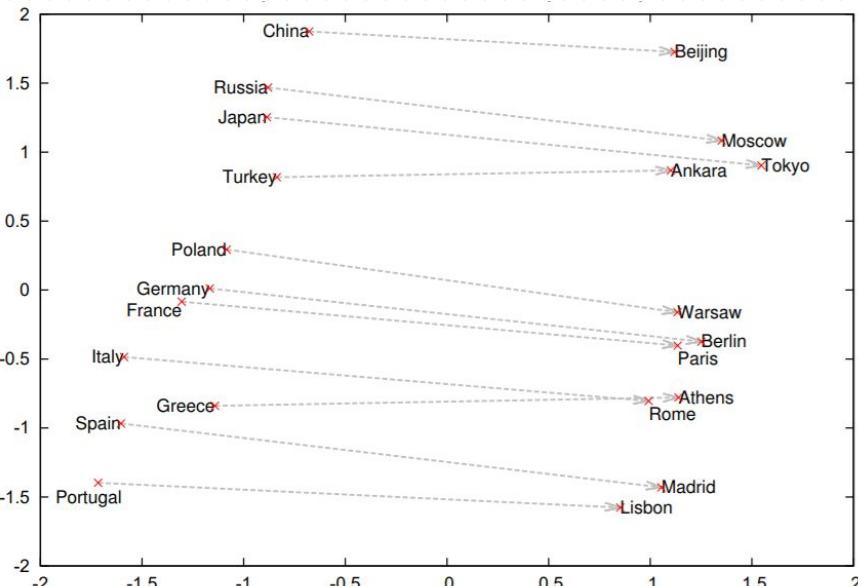
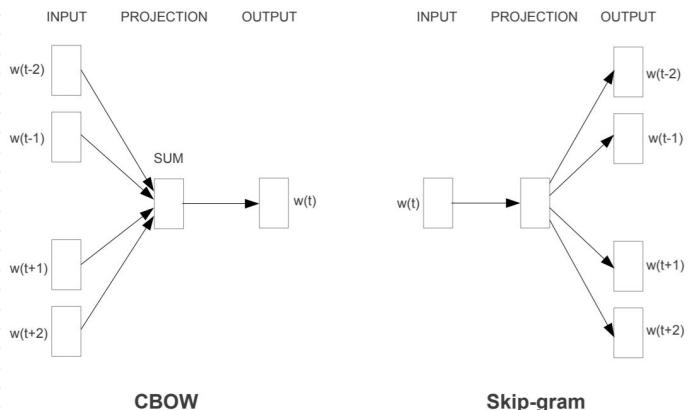
- ✗ Pode ser considerado uma forma de Aprendizado Não Supervisionado, onde os dados provém supervisão
- ✗ Em geral, treina-se o modelo para tentar “predizer os dados”
- ✗ É uma forma de Representation Learning que não necessita de dados rotulados



# Diferenciando de outros paradigmas

## Self-supervised

### X Exemplo: Word2Vec

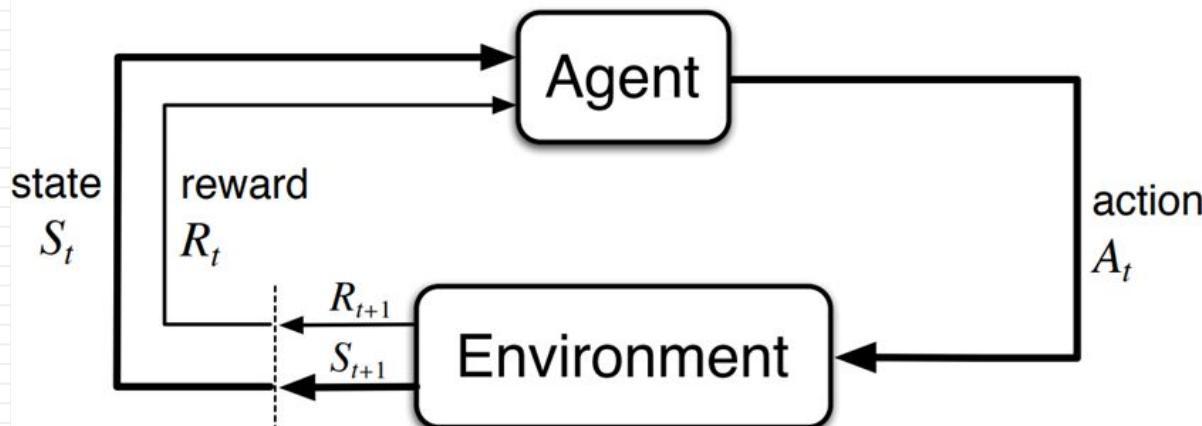


Fonte: Mikolov, Tomas, et al. "Distributed representations of words and phrases and their compositionality." Advances in neural information processing systems. 2013.



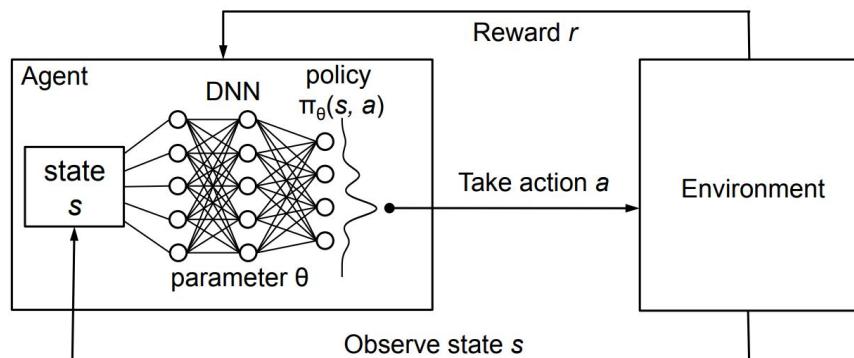
# Diferenciando de outros paradigmas

## Reinforcement Learning



# Diferenciando de outros paradigmas

## Reinforcement Learning (DQN)



Fonte: <https://www.youtube.com/watch?v=TmPfTpjtdgg>



# Aprendizado Não Supervisionado

Alguns exemplos onde o paradigma pode ser aplicado



# Segmentação de clientes

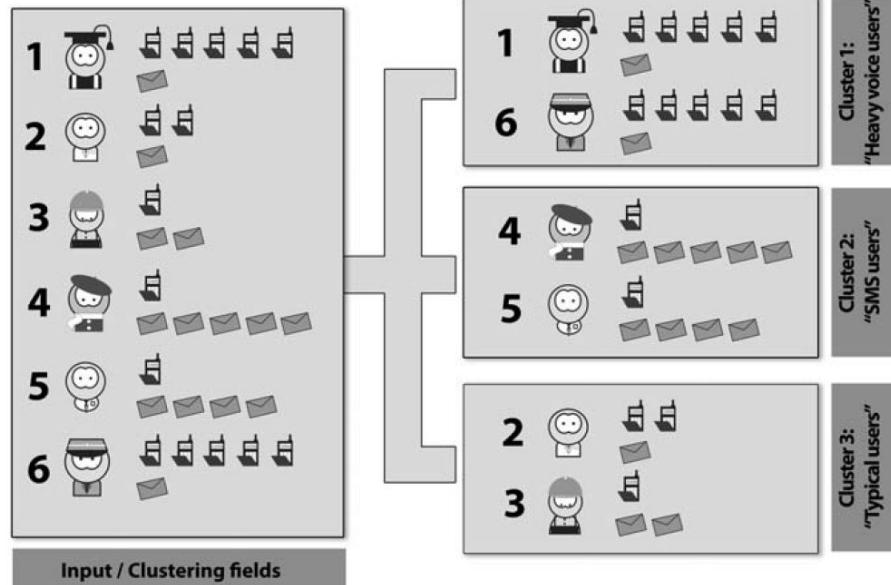
---

- ✖ Refere-se ao processo de dividir clientes em subgrupos distintos, significativos e homogêneos, com base em vários atributos e características;
- ✖ Tradicionalmente, empresas utilizam métodos de marketing baseados em dados demográficos e algumas outras informações;
- ✖ No entanto, para se obter vantagem competitiva deve-se utilizar técnicas mais adequadas para entender melhor os clientes.

Fonte: Tsipitsis, Konstantinos K., and Antonios Chorianopoulos. Data mining techniques in CRM: inside customer segmentation. John Wiley & Sons, 2011.



# Segmentação de clientes



## Pode ser útil para

- ✗ Melhor entendimento dos clientes
- ✗ Criação de novos produtos ou serviços
- ✗ Diferenciação no atendimento ao cliente
- ✗ Priorização de iniciativas de Marketing

Fonte: Tsipitsis, Konstantinos K., and Antonios Chorianopoulos. Data mining techniques in CRM: inside customer segmentation. John Wiley & Sons, 2011.



# Detecção de anomalias

---

- ✗ O objetivo é identificar instâncias anormais, que sejam diferentes da maior parte das outras instâncias;
- ✗ Instâncias anômalas são conhecidas como *outliers* (também chamado de Detecção de Outliers);
- ✗ Temos também o *Novelty Detection*
  - ✗ Procura detectar dados não observados anteriormente
  - ✗ Uma das principais diferenças: tais dados são normalmente incorporados no modelo original

Fonte: Tan, Pang-Ning. Introduction to data mining. Pearson Education India, 2007.



# Detecção de anomalias

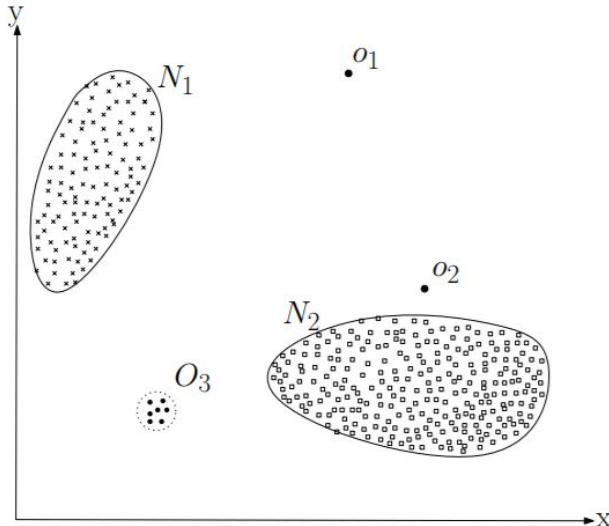


Fig. 1. A simple example of anomalies in a 2-dimensional data set.

Alguns exemplos incluem

- ✗ Detecção de fraudes em cartões de crédito
- ✗ Saúde pública
- ✗ Detecção de intrusão em sistemas de informação

Fonte: Chandola, Varun, Arindam Banerjee, and Vipin Kumar. "Anomaly detection: A survey." ACM computing surveys (CSUR) 41.3 (2009): 15.



# Regras de associação

---

- ✗ Técnica amplamente utilizada para analisar dados transacionais, com o objetivo de encontrar regras a partir de tais dados de acordo com uma medida de interesse;
- ✗ Não apenas regras, mas associações, correlações e padrões frequentes em grupos de itens;
- ✗ As regras encontradas geralmente possuem o formato SE/ENTÃO
  - ✗ **SE:** antecedente **ENTÃO:** consequente (efeito)
  - ✗ Exemplo: “Se um cliente compra carvão, ele tem 80% de chance de comprar carne”



# Regras de associação

- Exemplo clássico (*market basket transactions*)

ID da transação	Itens
1	{Bread, Milk}
2	{Bread, Diapers, Beer, Eggs}
3	{Milk, Diapers, Beer, Cola}
4	{Bread, Milk, Diapers, Beer}
5	{Bread, Milk, Diapers, Cola}

**Regra possível**

- $\{Diapers\} \rightarrow \{Beer\}$

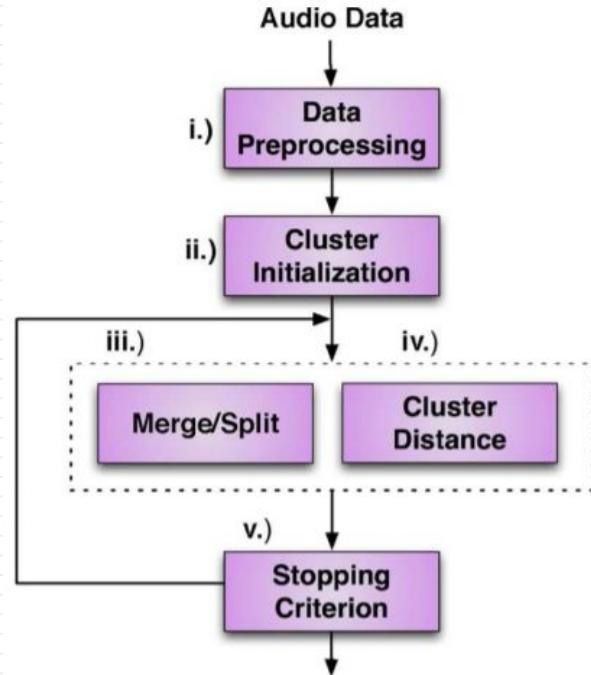
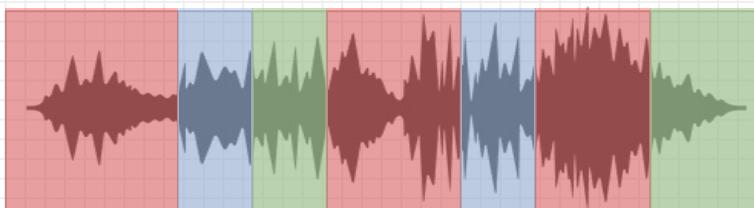
Sugere que existe forte relação entre a venda de fraldas e cerveja.

Fonte: Tan, Pang-Ning. Introduction to data mining. Pearson Education India, 2007.



# Speaker Diarization

- ✗ **Quem falou quando**
- ✗ Identificação não supervisionada dos locutores de um áudio e os intervalos nos quais cada locutor está falando



Fonte: Anguera, Xavier, et al. "Speaker diarization: A review of recent research." IEEE Transactions on Audio, Speech, and Language Processing 20.2 (2012): 356-370.



# Topic Modeling

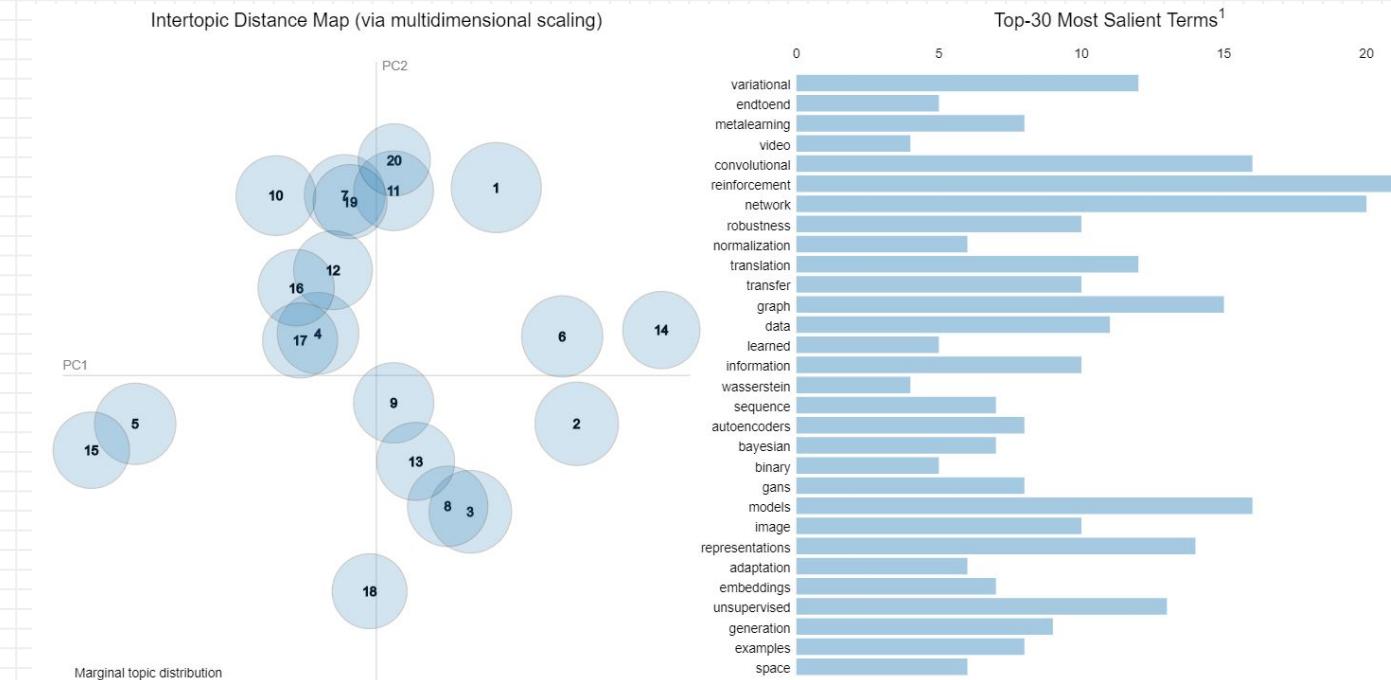
---

- ✗ Técnica para analisar uma grande coleção de documentos
  - ✗ Descobrir temas “escondidos” que estão na coleção
  - ✗ Rotular documentos de acordo com esses temas
  - ✗ Utilizar as anotações para organizar, resumir e buscar textos
- ✗ Um tópico é definido com base na co-ocorrência de palavras e inclui grupos de palavras que geralmente ocorrem em conjunto em diferentes documentos;
- ✗ Geralmente utilizado como uma etapa inicial antes de aplicar-se tarefas de mineração de texto.

Fonte: Barde, Bhagyashree Vyankatral, and Anant Madhavrao Bainwad. "An overview of topic modeling methods and tools." Intelligent Computing and Control Systems (ICICCS), 2017 International Conference on. IEEE, 2017.



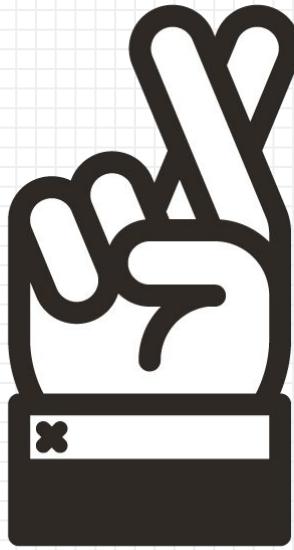
# Topic Model – Exemplo



Fonte: ICLR 2019 papers



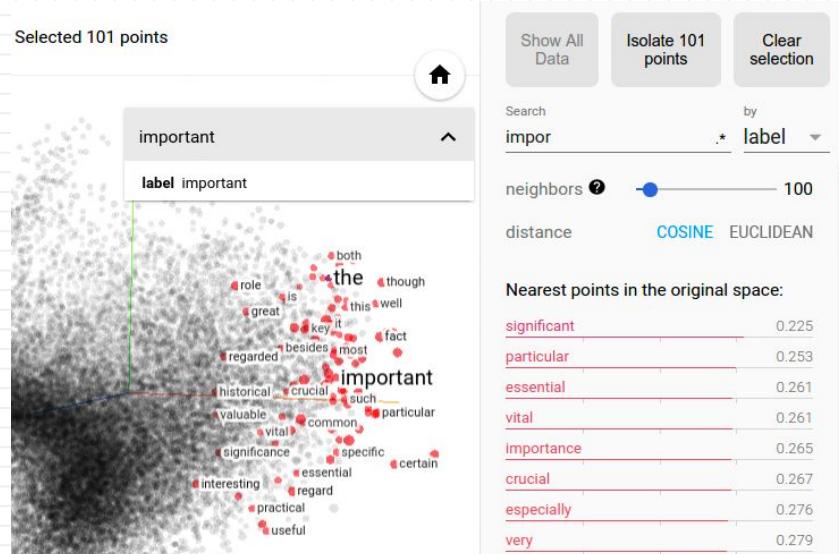
# DEMO TIME



# Redução de Dimensionalidade – Exemplos

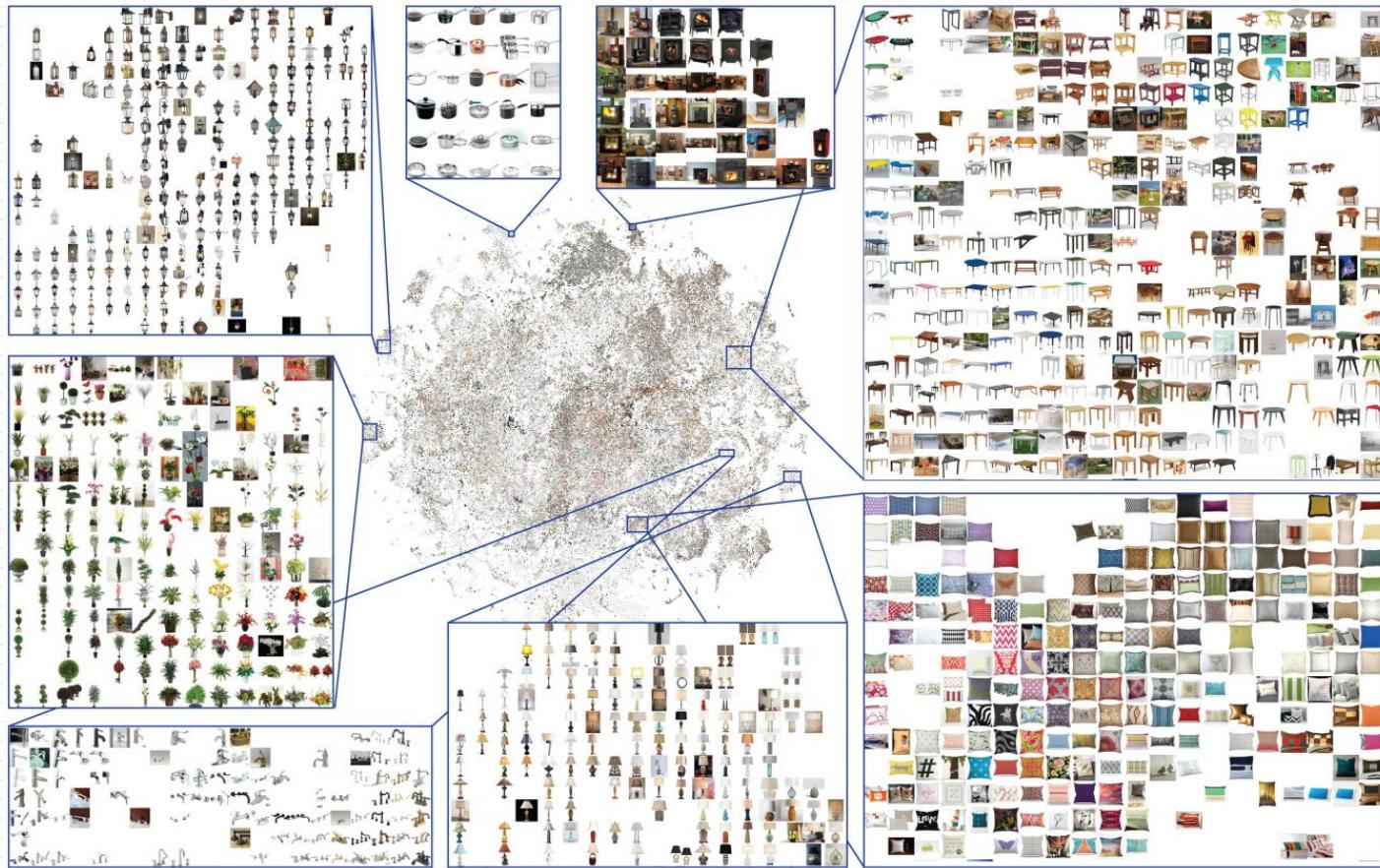


Fonte: <https://cs.stanford.edu/people/karpathy/cnnembed/>



Fonte: <https://www.tensorflow.org/guide/embedding>

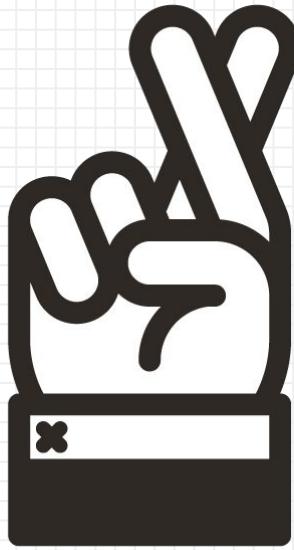




Fonte: Bell, Sean, and Kavita Bala. "Learning visual similarity for product design with convolutional neural networks." ACM Transactions on Graphics (TOG) 34.4 (2015): 98.



# DEMO TIME



# Aprendizado Não Supervisionado

Alguns exemplos do estado-da-arte



# Image-to-Image Translation

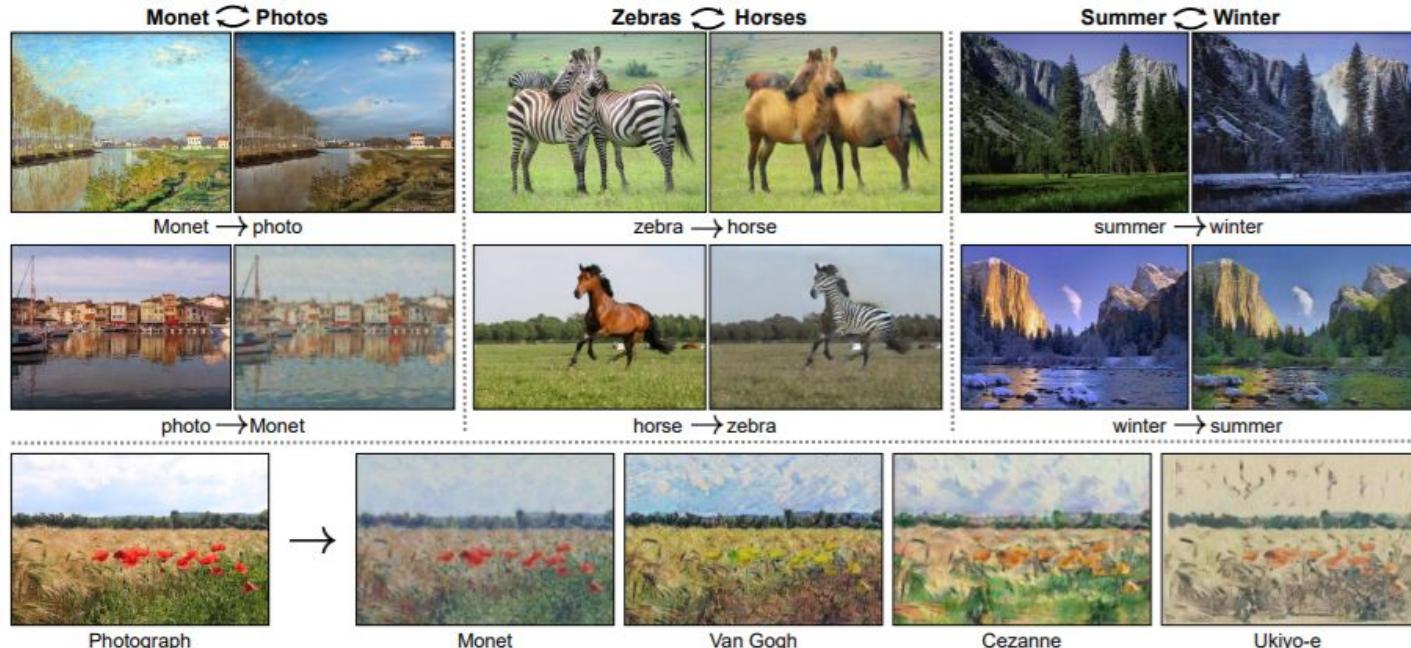
---

- ✗ Aprender o mapeamento de uma imagem de entrada para uma imagem de saída;
- ✗ Utiliza a imagem de entrada como uma condição para a geração da imagem de saída;
- ✗ Baseado em *Generative Adversarial Networks* (GANs)
  - ✗ Veremos mais adiante na aula o que elas são!

Fonte: Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." arXiv preprint (2017).



# Image-to-Image Translation



Fonte: Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." arXiv preprint[2017].



# Video-to-Video Synthesis

---

- ✗ Aprender o mapeamento de um vídeo de entrada (geralmente simplificado, como uma segmentação) para um vídeo de saída foto realístico;
- ✗ Primeiro trabalho que conseguiu implementar o conceito de *image-to-image translation* para vídeos, respeitando o aspecto temporal;
- ✗ Também é baseado em GANs

Fonte: Wang, Ting-Chun, et al. "Video-to-Video Synthesis." arXiv preprint arXiv:1808.06601 (2018).



# Video-to-Video Synthesis



Fonte: Wang, Ting-Chun, et al. "Video-to-Video Synthesis." arXiv preprint arXiv:1808.06601 (2018).

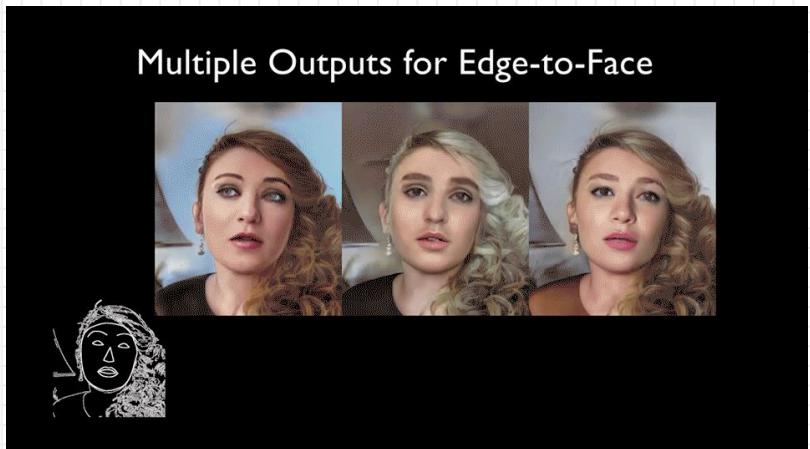


# Video-to-Video Synthesis

Edge-to-Face Results



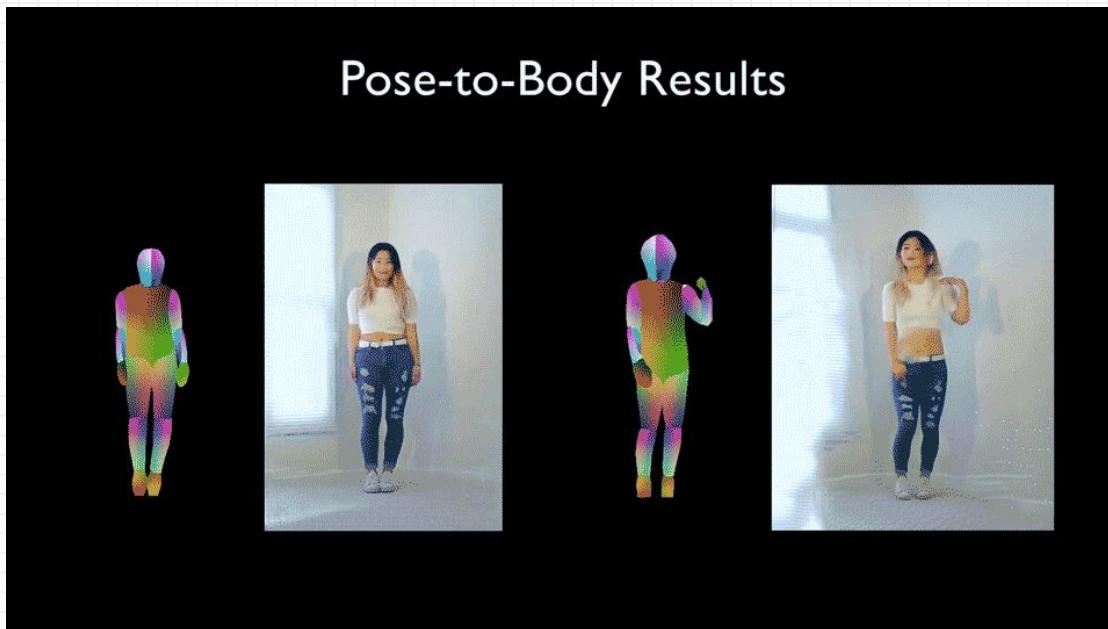
Multiple Outputs for Edge-to-Face



Fonte: Wang, Ting-Chun, et al. "Video-to-Video Synthesis." arXiv preprint arXiv:1808.06601 (2018).



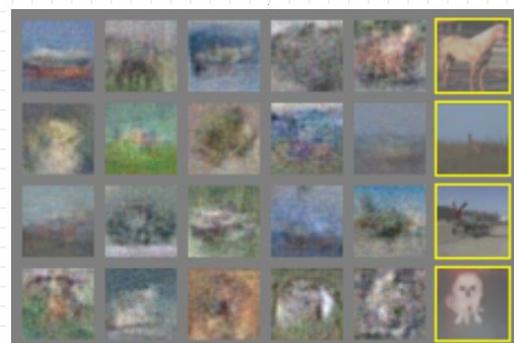
# Video-to-Video Synthesis



Fonte: Wang, Ting-Chun, et al. "Video-to-Video Synthesis." arXiv preprint arXiv:1808.06601 (2018).



# Image Generation (no início)



Fonte: Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.



# Image Generation



Figure 5:  $1024 \times 1024$  images generated using the CELEBA-HQ dataset. See Appendix F for a larger set of results, and the accompanying video for latent space interpolations.

Our ( $256 \times 256$ )

Fonte: Karras, Tero, et al. "Progressive growing of gans for improved quality, stability, and variation." arXiv preprint arXiv:1710.10196 (2017).



# Caso de uso específico: imagens médicas

- ✗ Um survey de 2017 encontrou mais de 300 artigos científicos que citavam Deep Learning e Medical Image Analysis;
- ✗ Contudo, grandes quantidades de dados rotulados e com qualidade são complexos de se obter
  - ✗ São necessários especialistas treinados e tempo
  - ✗ *Transfer Learning* poderia ajudar
    - De igual forma, precisa de grande quantidade de dados

Fonte: Perone, Christian S., et al. "Unsupervised domain adaptation for medical imaging segmentation with self-ensembling." arXiv preprint arXiv:1811.06042 (2018).



# Caso de uso específico: imagens médicas

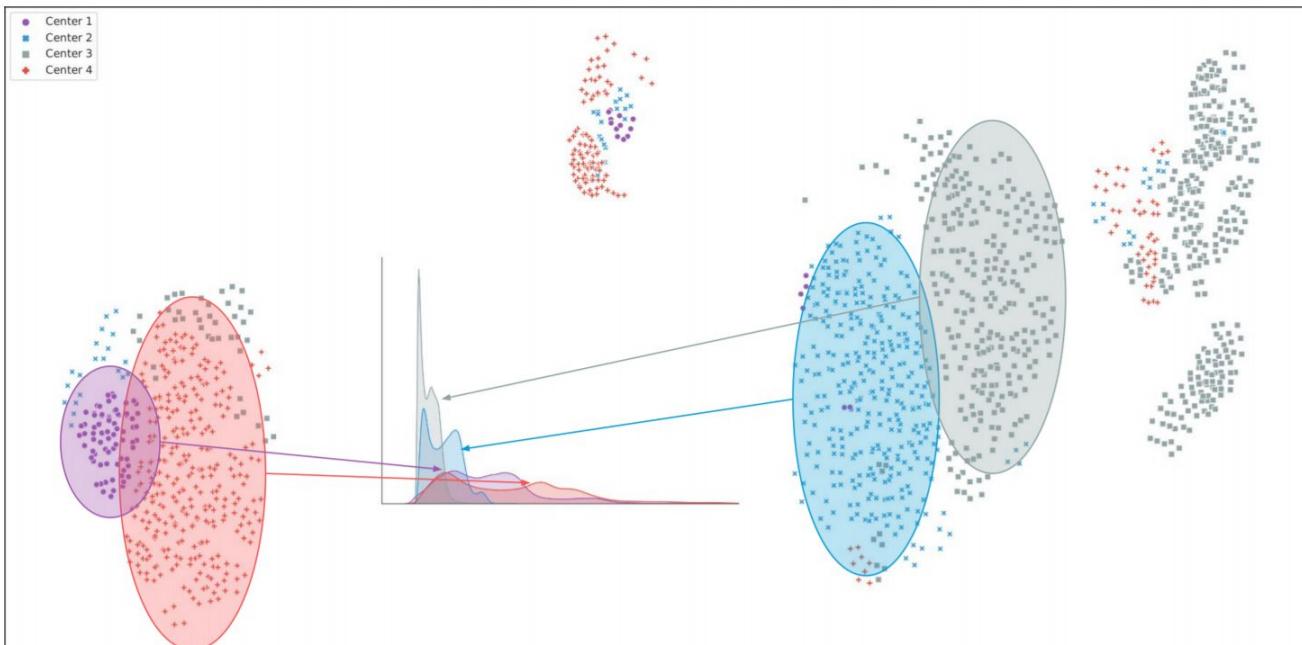


Figure 9. Expanded visualization for t-SNE from the adaptation scenario in Figure 8b. The chart in the middle represents the pixel distribution from each center. It can be observed how similar distributions tend to form clusters on the prediction space.

Fonte: Perone, Christian S., et al. "Unsupervised domain adaptation for medical imaging segmentation with self-ensembling." arXiv preprint arXiv:1811.06042 (2018).



# Como o Aprendizado Não Supervisionado ajudou Deep Learning



# Pré-treino Não Supervisionado

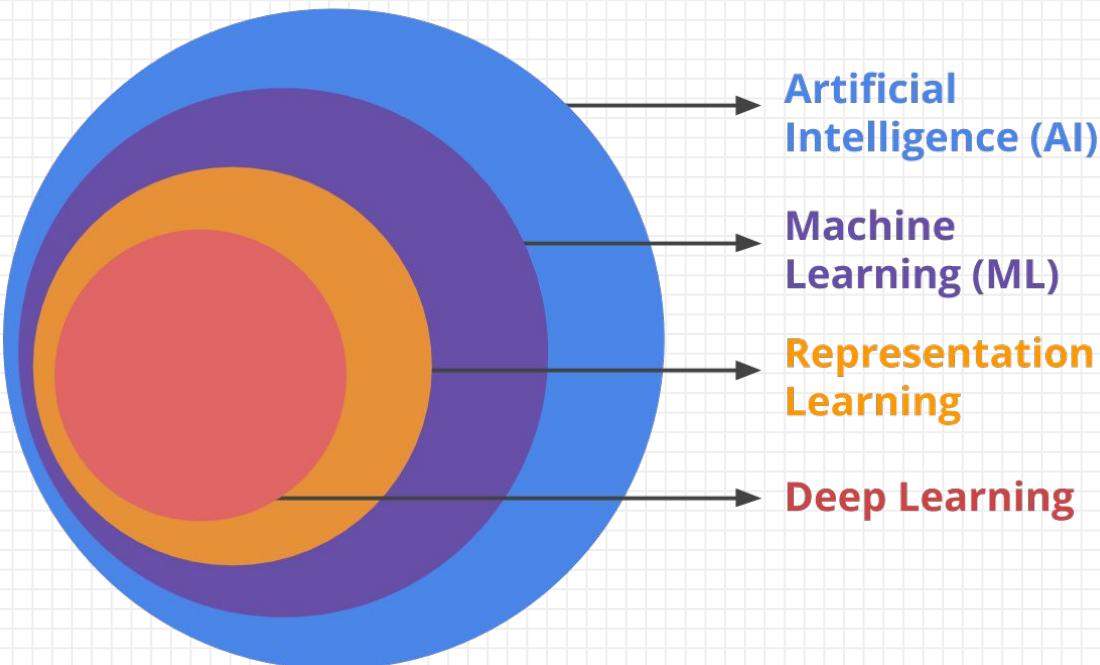
---

- ✗ Durante alguns anos, foram efetuadas diferentes tentativas para treinar redes neurais profundas (com mais de 1 ou 2 camadas escondidas)
  - ✗ Tentativas não eram bem sucedidas (*vanishing gradient*)
- ✗ Um *breakthrough* aconteceu em 2006:
  - ✗ Aprendizado não supervisionado como pré-treino;
  - ✗ Aprendizado supervisionado para realizar *fine-tuning*
- ✗ Atualmente, pré-treino não supervisionado é praticamente não utilizado

Fonte: Erhan, Dumitru, et al. "Why does unsupervised pre-training help deep learning?." Journal of Machine Learning Research 11.Feb (2010): 625-660..



# Deep Learning

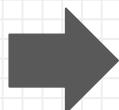


Fonte: Deep Learning (Goodfellow, Bengio, Courville)

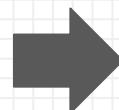


# Deep Learning

“Tradicional”

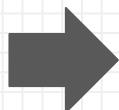


Extrator de  
features  
*hand-crafted*

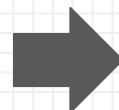


Classificador

Deep Learning



Extrator de  
features treinado



Classificador



# Interpretabilidade

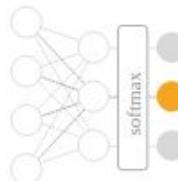
Different **optimization objectives** show what different parts of a network are looking for.

n layer index

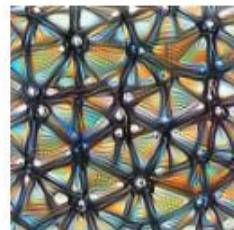
x, y spatial position

z channel index

k class index



Neuron  
 $\text{layer}_n[x, y, z]$



Channel  
 $\text{layer}_n[:, :, z]$



Layer/DeepDream  
 $\text{layer}_n[:, :, :]^2$



Class Logits  
 $\text{pre\_softmax}[k]$

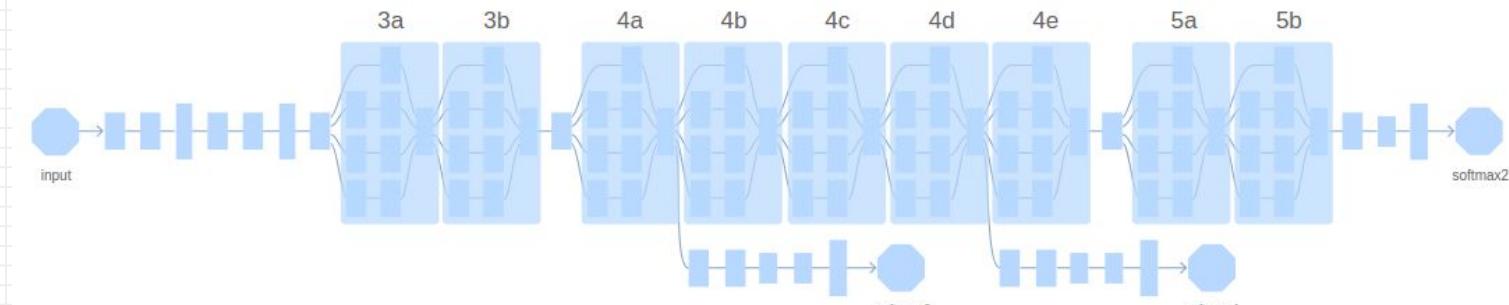


Class Probability  
 $\text{softmax}[k]$

Fonte: Feature Visualization - <https://distill.pub/2017/feature-visualization/>



# Interpretabilidade



## Layer 4c



Palm trees



Wheels



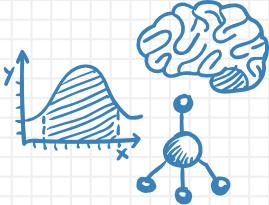
Dogs on leash



Houses

Fonte: Feature Visualization - <https://distill.pub/2017/feature-visualization/>





**“Human and animal learning is largely unsupervised: we discover the structure of the world by observing it, not by being told the name of every object.”**

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521.7553 (2015): 436.



# Aprendizado Não Supervisionado



# Considerações finais

---

- ✗ Aprendizado Não Supervisionado visa encontrar estruturas/padrões/regularidades nos dados;
- ✗ Existem mais aplicações e pesquisas relacionadas à Aprendizado Supervisionado do que Aprendizado Não Supervisionado
  - ✗ Combinação das duas abordagens possui grande potencial
- ✗ Aprendizado Não Supervisionado é muito mais do que apenas Agrupamento de Dados :)



# Vamos iniciar nossa jornada!

