

# SEARCHING FOR HIDDEN PATTERNS OF EMISSION LINES IN GALAXIES WITH DEEP LEARNING

---

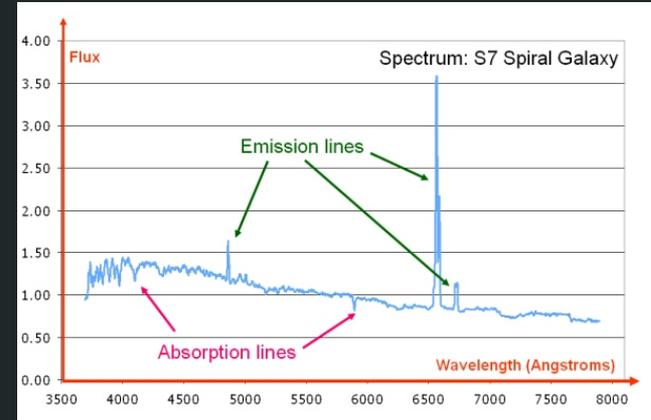
Patrick Schubert de Souza Fuchs  
Ciência da Computação - UFRJ

Clécio Roque de Bom, PhD

# O que são galáxias de linhas de emissão?

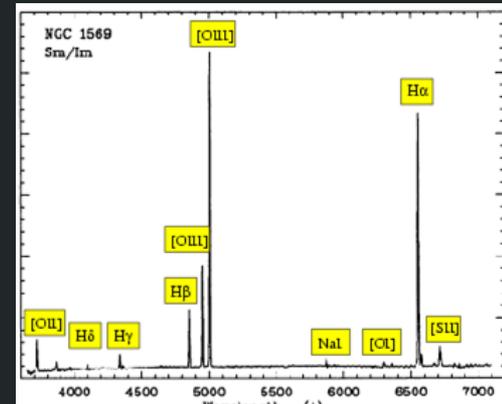
São galáxias caracterizadas pela radiação eletromagnética que emitem, como as galáxias de núcleo ativo e as galáxias com formação de estrelas.

A intensidade dessa radiação em determinada frequência está associada a presença de determinado elemento químico, o que pode ser usado para classificar o astro.



Porquê nos interessa a aplicação das técnicas de aprendizado profundo em linhas de emissão?

- .Crescimento do número de objetos com linhas de emissão extragalácticos e com espectro de alta resolução (V. Stampoulis et al).
- .Alta complexidade de tempo da análise de dados.
- .Inconsistência nos esquemas de classificação: ~8 % das galáxias são classificadas em classes diferentes em distintos diagramas.



## Como é feita a classificação de objetos de linhas de emissão atualmente?

Plots entre log de razões da largura equivalente das linhas de emissão, tipicamente esquemas WHAN e BPT. No gráfico ao lado foram usadas 4 razões, o que limita a quantidade de informação que poderia estar sendo usada na análise devido a complexidade dos algoritmos utilizados.

$$\begin{bmatrix} \text{Cov}(X, X) & \text{Cov}(Y, X) \\ \text{Cov}(X, Y) & \text{Cov}(Y, Y) \end{bmatrix}$$

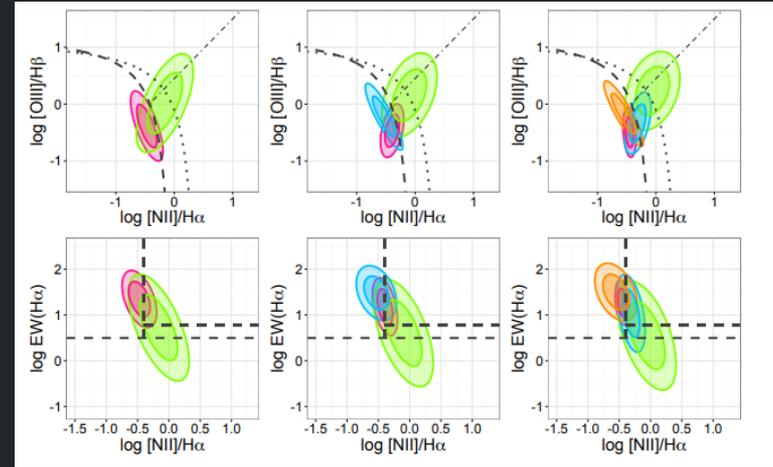


Imagem: De Souza et al (2017) arXiv: 1703.07607

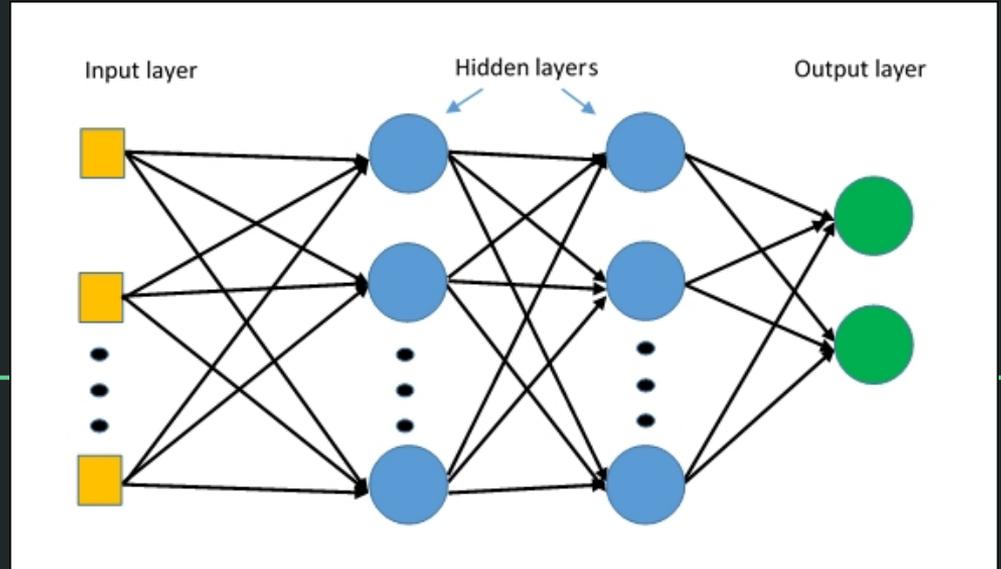
Como a abordagem em Deep Learning nos possibilita ter resultados mais robustos?

Os algoritmos de Deep Learning são pouco sensíveis à dimensão das entradas. O que significa que podemos analisar dados mais complexos com quase o mesmo número de parâmetros (Big Data).

Hbeta	OIII.4959	OIII.5007	NII.6548	Halpha	NII.6584	SII.6716	SII.6731
1.173	0.319	0.955	0.870	5.540000	2.599	1.668	1.627
6.589	0.823	2.468	3.074	30.132999	9.190	6.520	4.312
3.232	0.386	1.159	1.792	14.688000	5.356	2.828	1.920
5.671	1.172	3.514	2.509	24.448999	7.499	4.689	3.506
4.820	1.437	4.309	1.311	19.671000	3.920	4.699	4.405

Como a abordagem em Deep Learning nos possibilita ter resultados mais robustos?

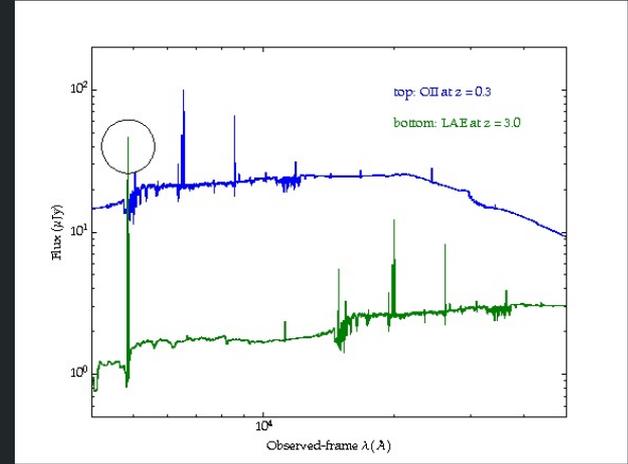
Em uma rede neural, a adição de mais uma feature na camada de entrada, resulta em um aumento de parâmetros igual ao número de nós na camada seguinte.



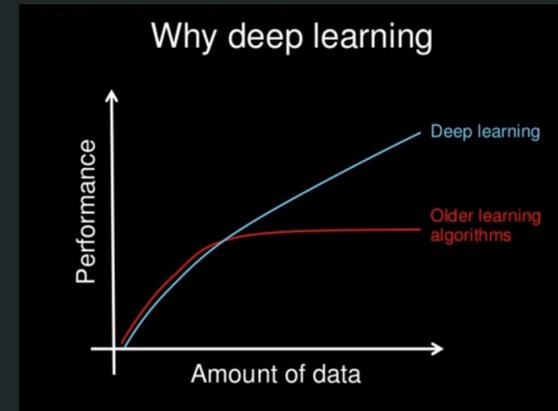
Como a abordagem em Deep Learning nos possibilita ter resultados mais robustos?

Como em toda análise de processamento de sinais, podemos ter problemas com ruídos. No caso das linhas de emissão, o desvio para o vermelho pode deslocar algumas linhas e fazer-nos acreditar que há presença de determinado elemento enquanto deveria haver outro.

Algoritmos de Deep Learning também são menos sensíveis a ruídos



Apesar de todas as competências anteriormente citadas, a principal qualidade dos algoritmos de deep learning é a capacidade de performar cada vez melhor em relação ao aumento do número de dados.



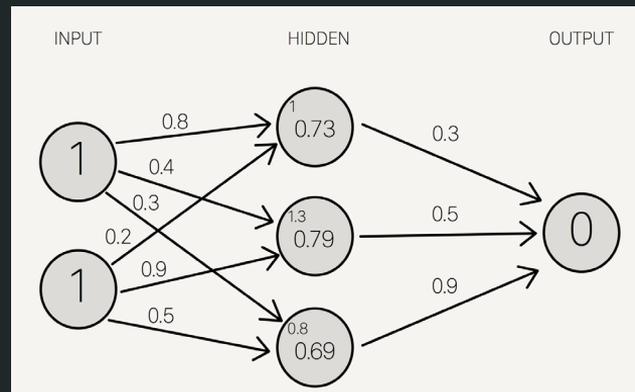
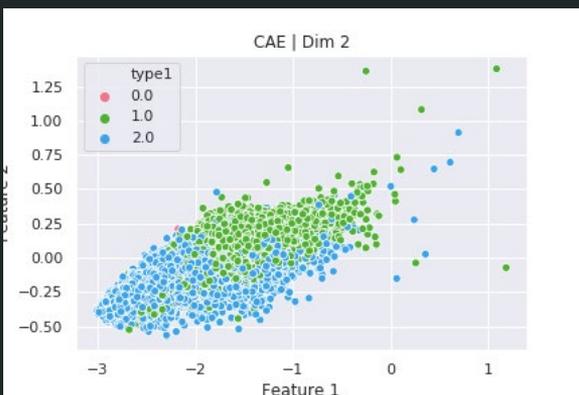
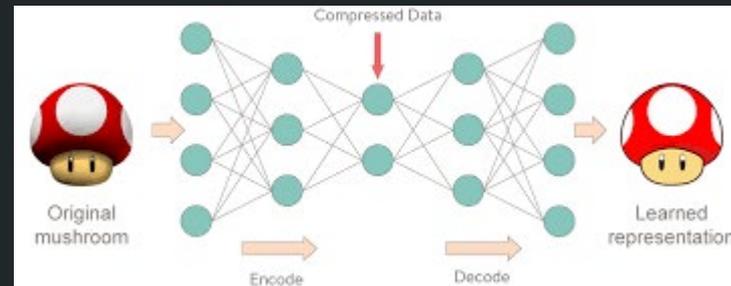
## Aprendizado não supervisionado:

Como redes neurais aprendem a separar dados?

- AutoEncoders

A principal característica do autoencoder é a capacidade de comprimir e reconstruir o dado de entrada. Quando essa reconstrução atinge baixos valores de erro, para um grande conjunto de dados diferentes, de graça conseguimos uma representação inequívoca em dimensão menor.

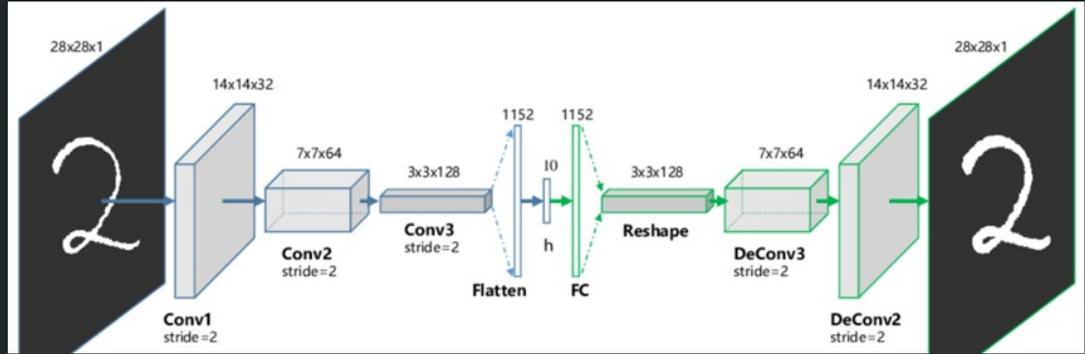
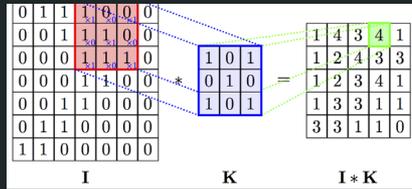
Logo, conseguiremos reconstruir dados diferentes fixando os valores dos pesos da rede e com isso, dados diferentes terão que ter representações diferentes no espaço.



## -Convolutional Autoencoder:

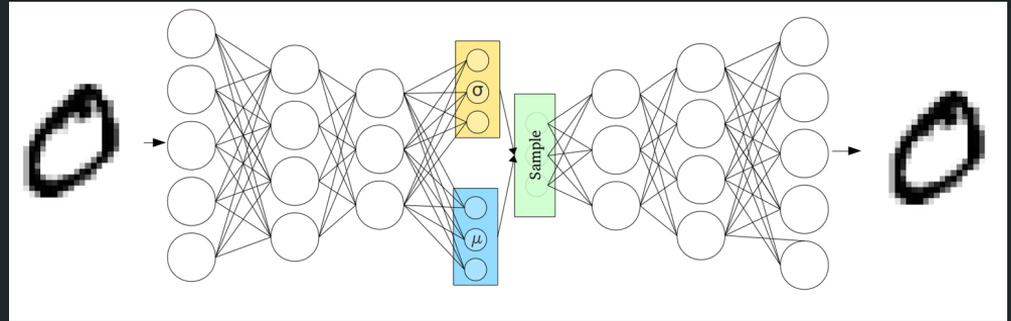
O Convolutional Autoencoder se diferencia do autoencoder comum por possuir camadas convolucionais.

Essas camadas possuem a qualidade de extrair features mais significativas dos dados, eliminando problemas de localidade e descartando features não impactantes na representação do dado.



## -Variational Autoencoder:

O Variational Autoencoder possui o mesmo propósito do autoencoder, porém esta arquitetura gera uma distribuição de probabilidade dos dados, otimizando um vetor de médias e um vetor de desvio-padrão que são usados para gerar novas representações dos dados



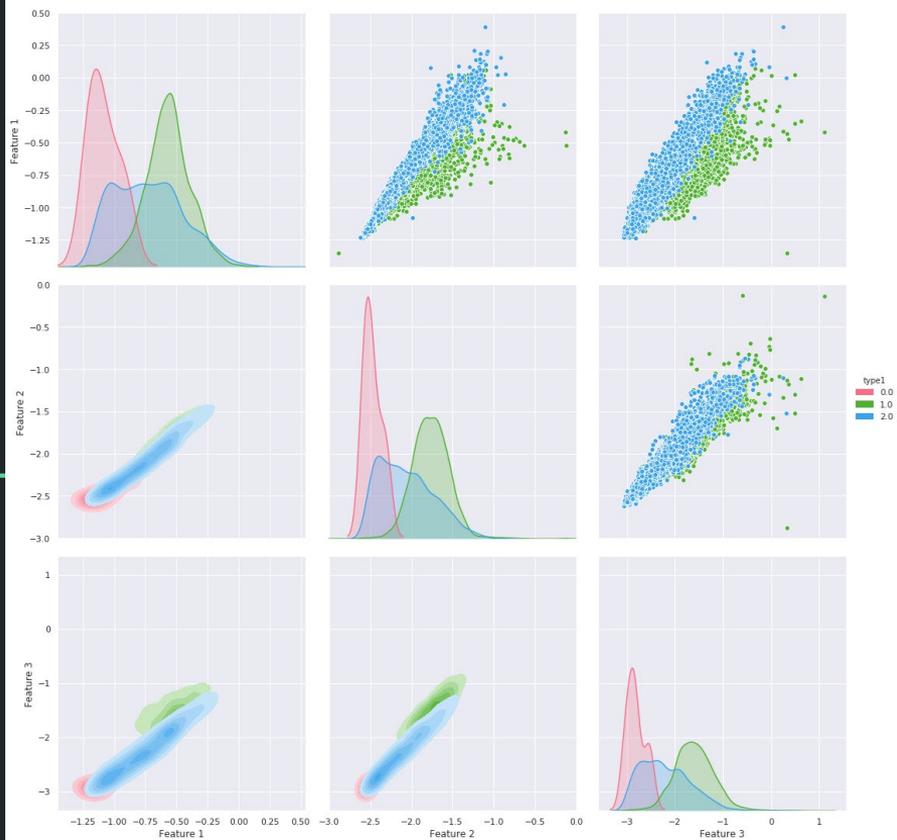
Conjunto de dados utilizado:

Utilizamos 8 linhas de emissão na análise de dados para 113281 mil galáxias em esquemas de classificação Whan e Type 1 (BPT).

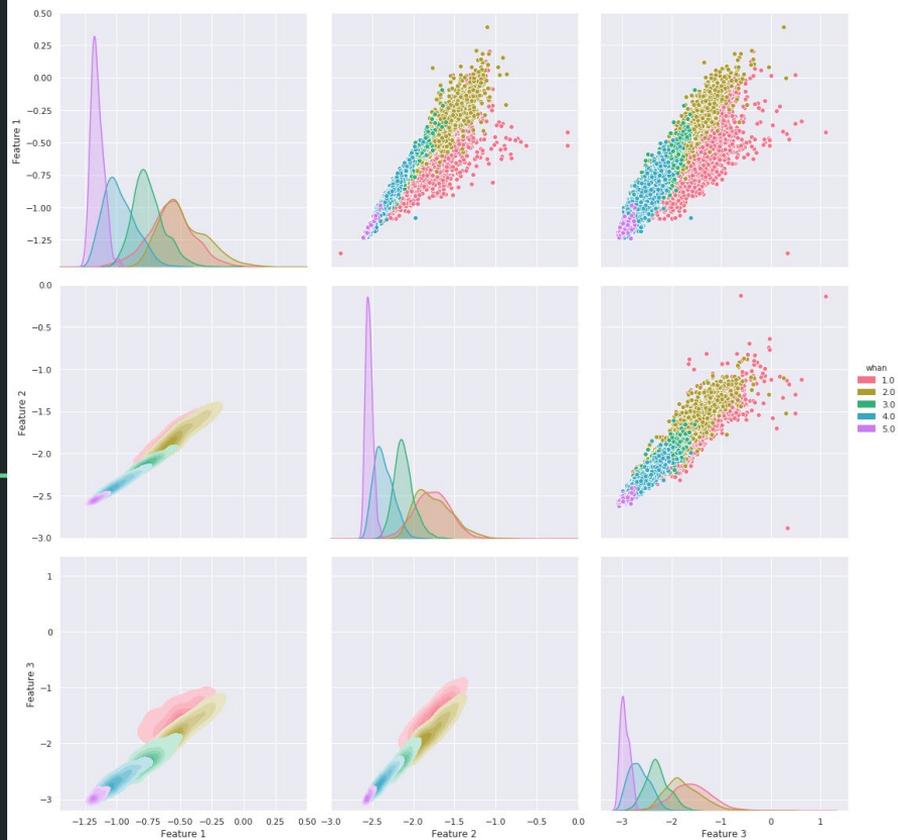
Hbeta	OIII.4959	OIII.5007	NII.6548	Halpha	NII.6584	SII.6716	SII.6731
1.173	0.319	0.955	0.870	5.540000	2.599	1.668	1.627
6.589	0.823	2.468	3.074	30.132999	9.190	6.520	4.312
3.232	0.386	1.159	1.792	14.688000	5.356	2.828	1.920
5.671	1.172	3.514	2.509	24.448999	7.499	4.689	3.506
4.820	1.437	4.309	1.311	19.671000	3.920	4.699	4.405

# Resultados dos espaços latentes: Convolutional Autoencoder

[CAE][Dim 3]

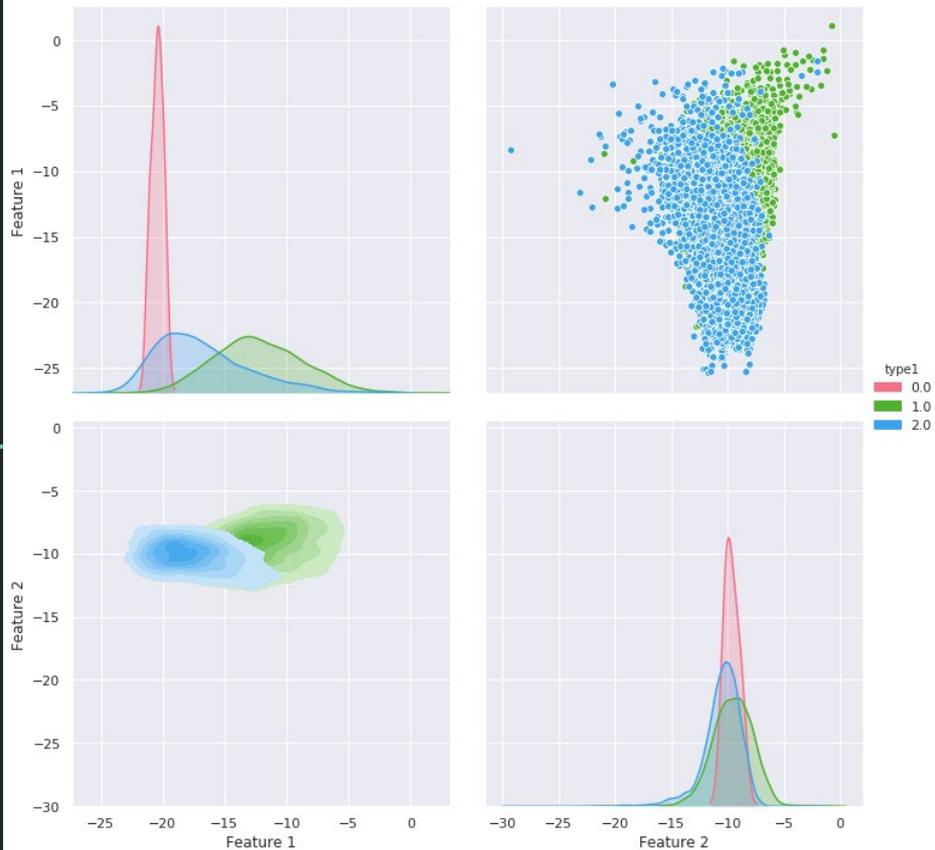


[CAE][Dim 3]

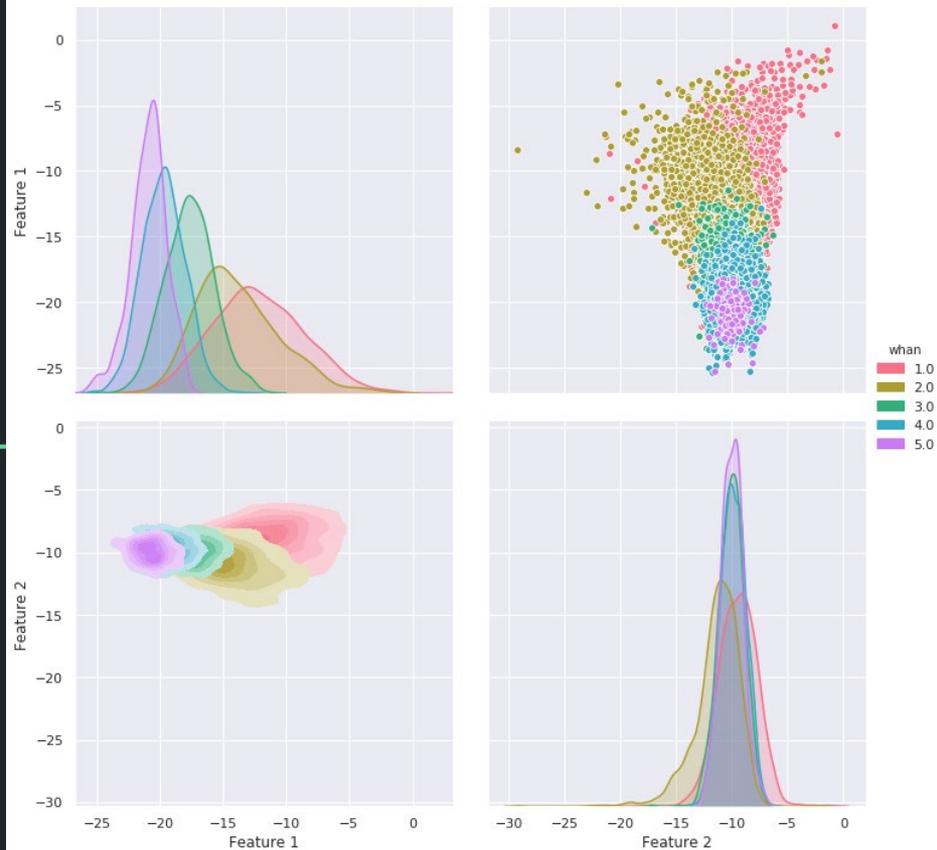


# Resultados dos espaços latentes: Convolutional Variational Autoencoder

[CVAE][Dim 2 | STD 1.0]

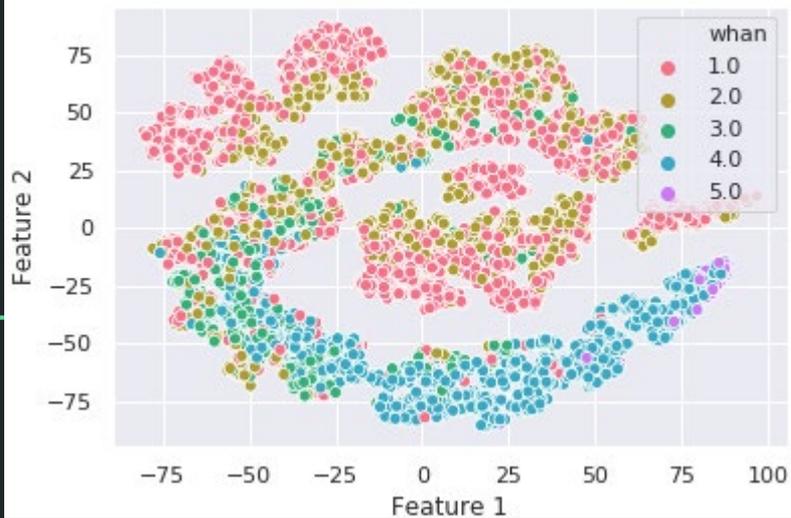


[CVAE][Dim 2 | STD 1.0]

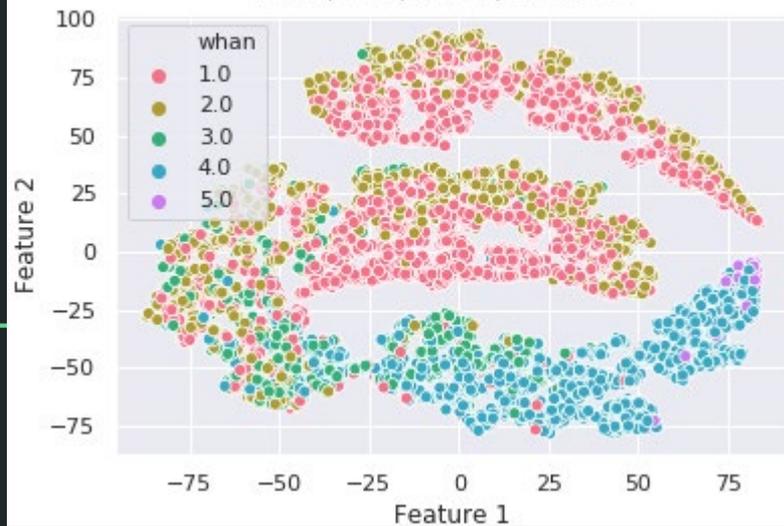


## Resultados em TSNE:

T-SNE | AE | Dim 3



T-SNE | VAE | Dim 4 | STD 1e-13

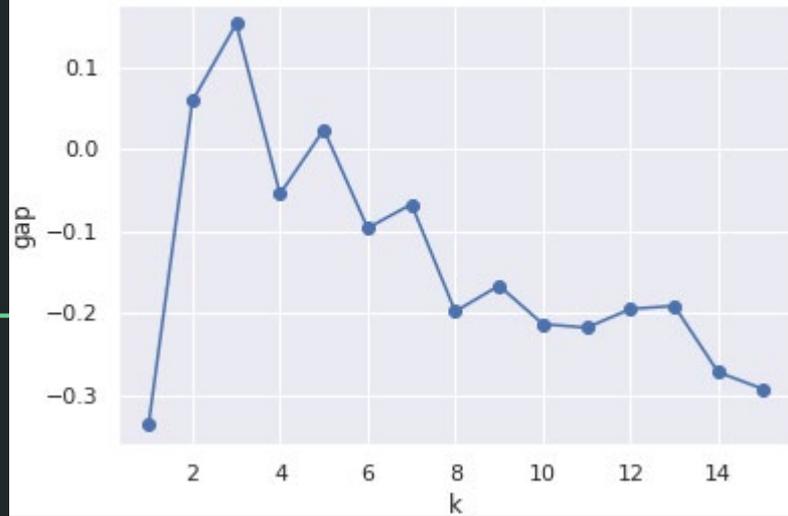


Quantas classes (clusters) conseguimos extrair dos dados?

[VAE][Dim 3 | STD 1]



[AE][Dim 2]



## Considerações finais:

- . Os esquemas de classificações não são perfeitos e geram inferências muitas vezes contraditórias.
- . O número de classes não é inequívoco, assim como as suas fronteiras.

. Otimismo quanto a possibilidade de definir um conjunto de classes e suas fronteiras de forma objetiva.

---

## Passos futuros:

- . Otimização de hiperparâmetros
- . Conjuntos de dados maiores e mais complexos
- . Treinos Adversariais

