# O que é “Deep Learning”

# & como posso usar na Astronomia?

—

André ZAMORANO VITORELLI

CosmoStat

cea

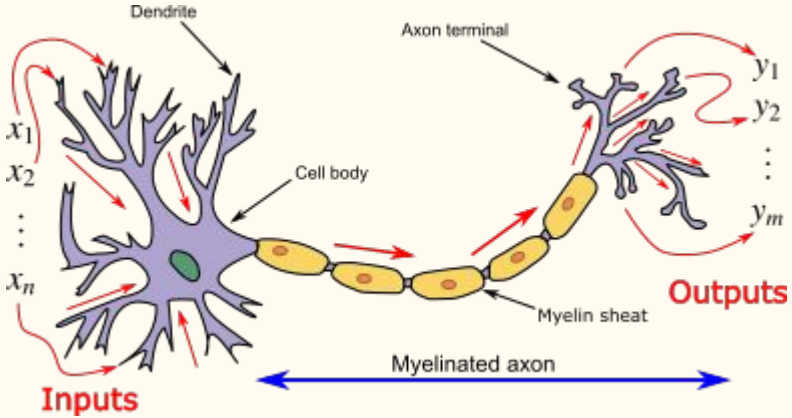
# O que são redes neurais?
# O que é “Deep Learning”?

# O conceito

# Neurônios biológicos

# Neurônios artificiais

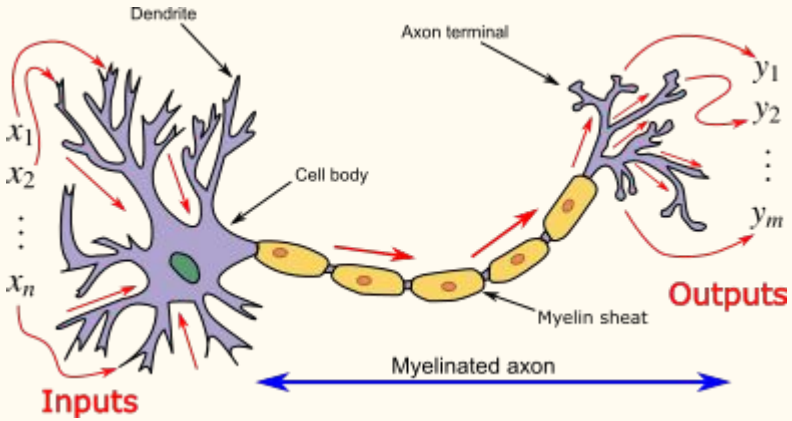
# Neurônios biológicos



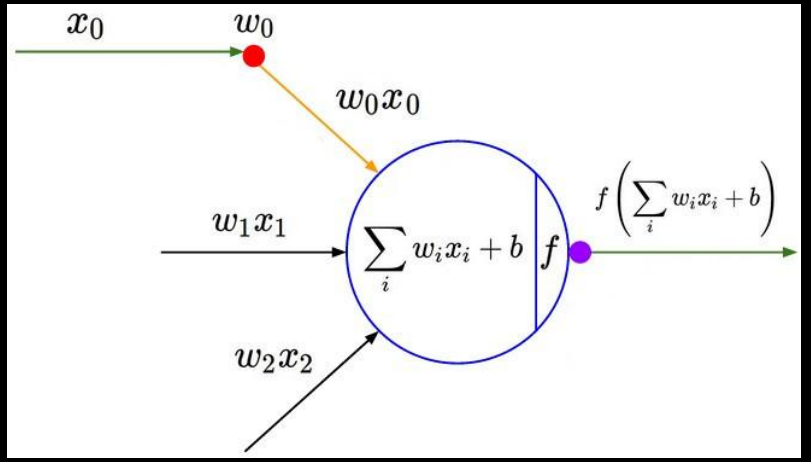

Wikipedia

# Neurônios artificiais

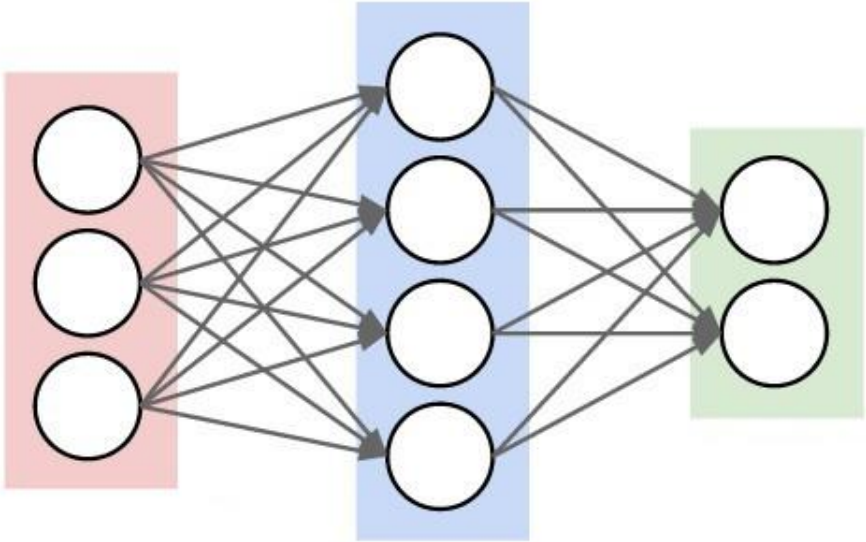
# Neurônios biológicos



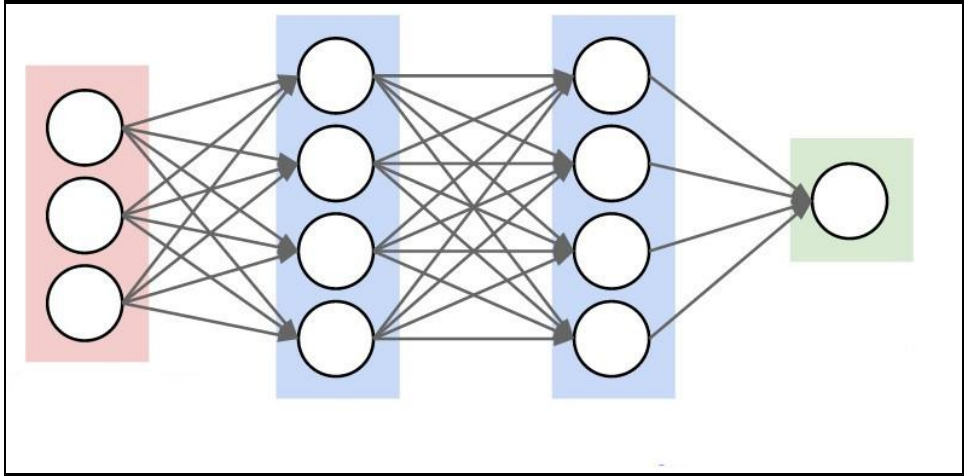

# Neurônios artificiais



# Neural Network

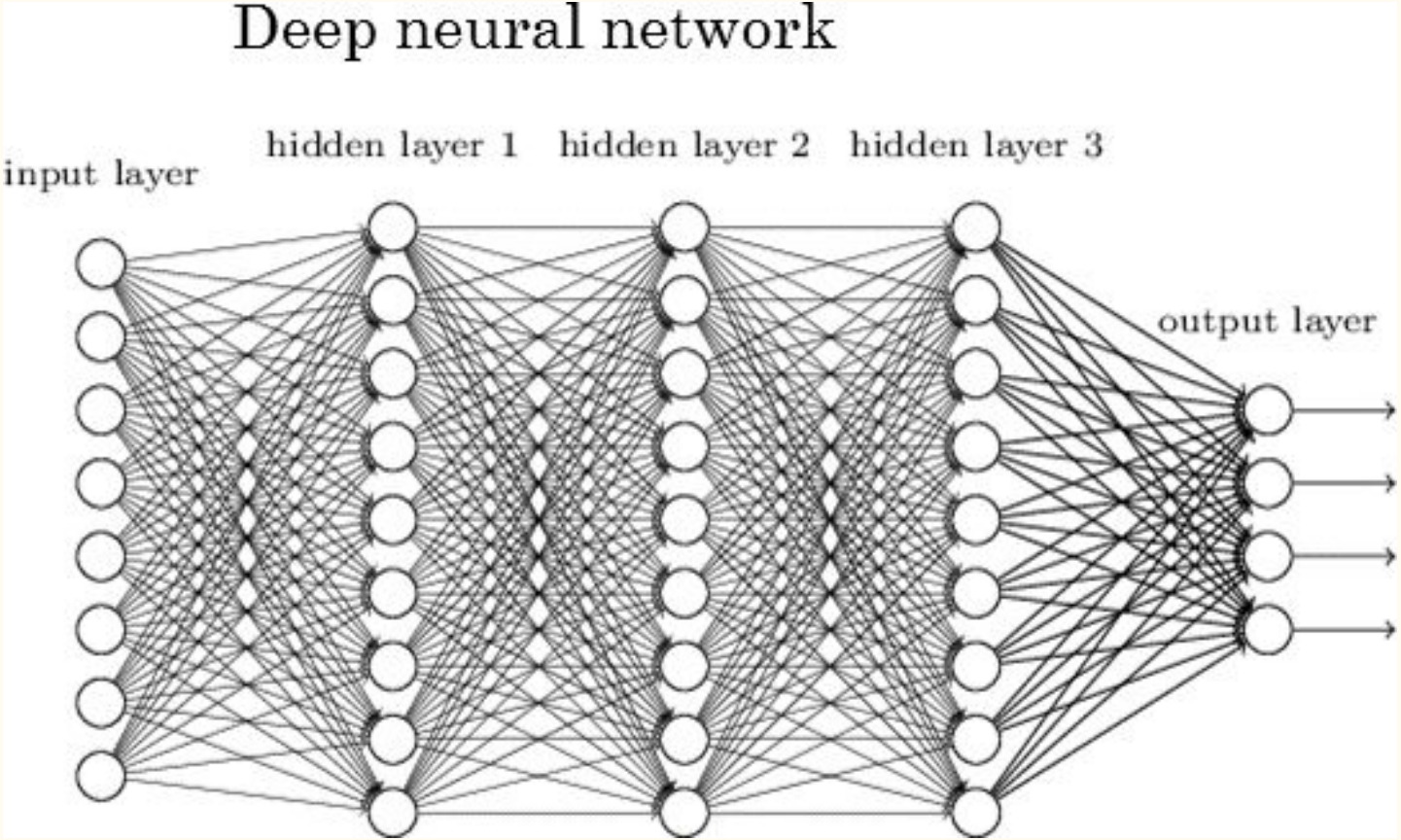


# Deep Neural Networks

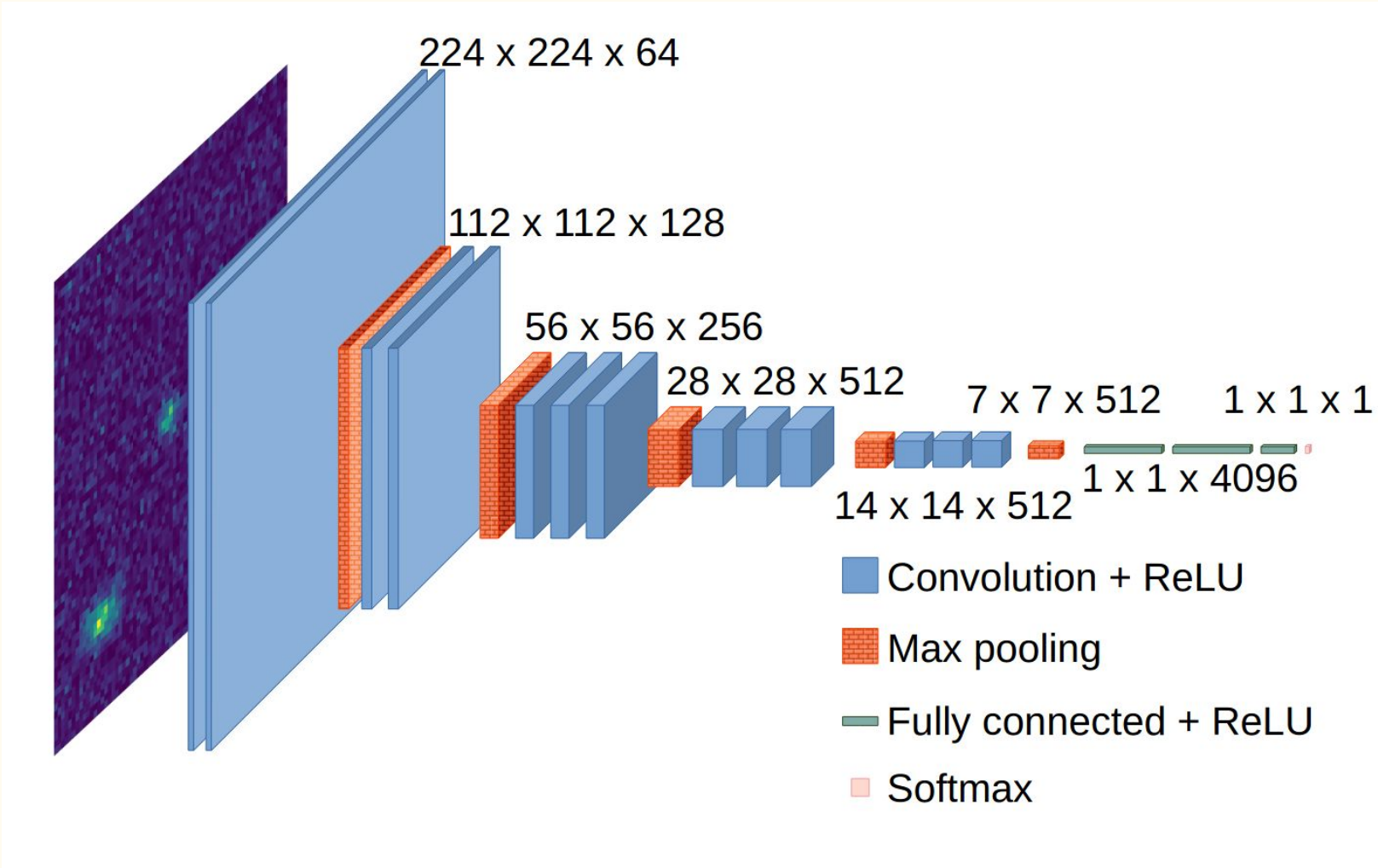


http://towardsdatascience.com

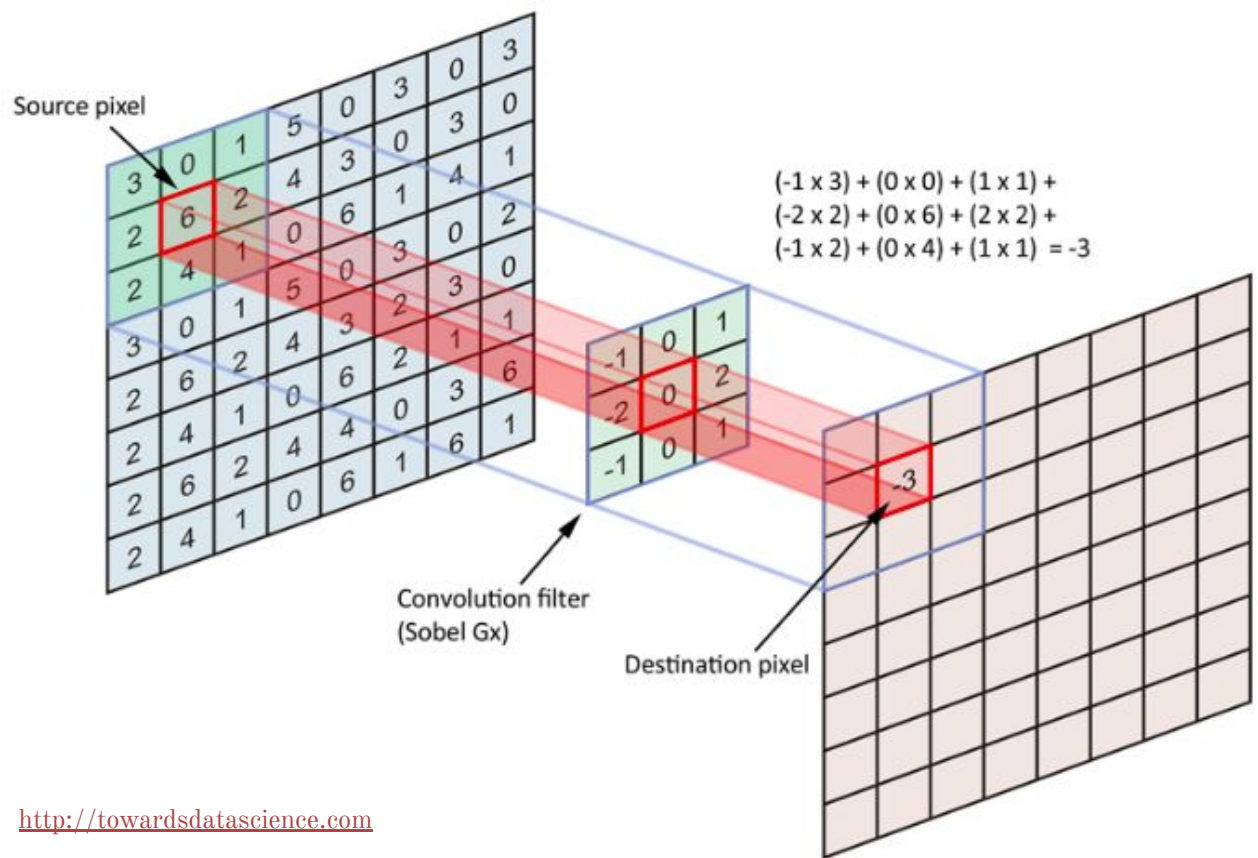
# Convolutional Neural Networks




http://towardsdatascience.com

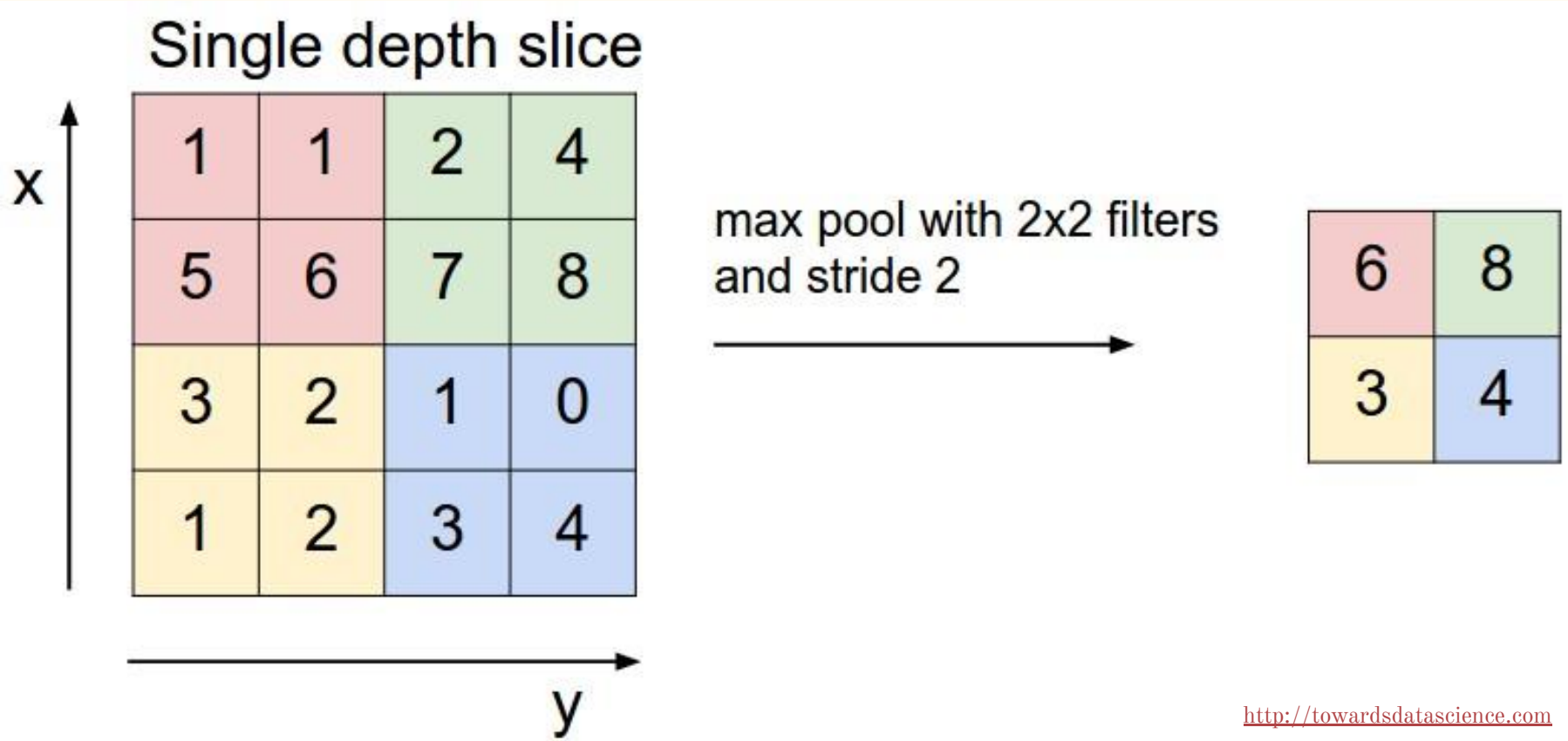

224 x 224 x 64
112 x 112 x 128
56 x 56 x 256
28 x 28 x 512
14 x 14 x 512
7 x 7 x 512
1 x 1 x 4096
1 x 1 x 1
Convolution + ReLU
Max pooling
Fully connected + ReLU
Softmax

# O que é um filtro de convolução?

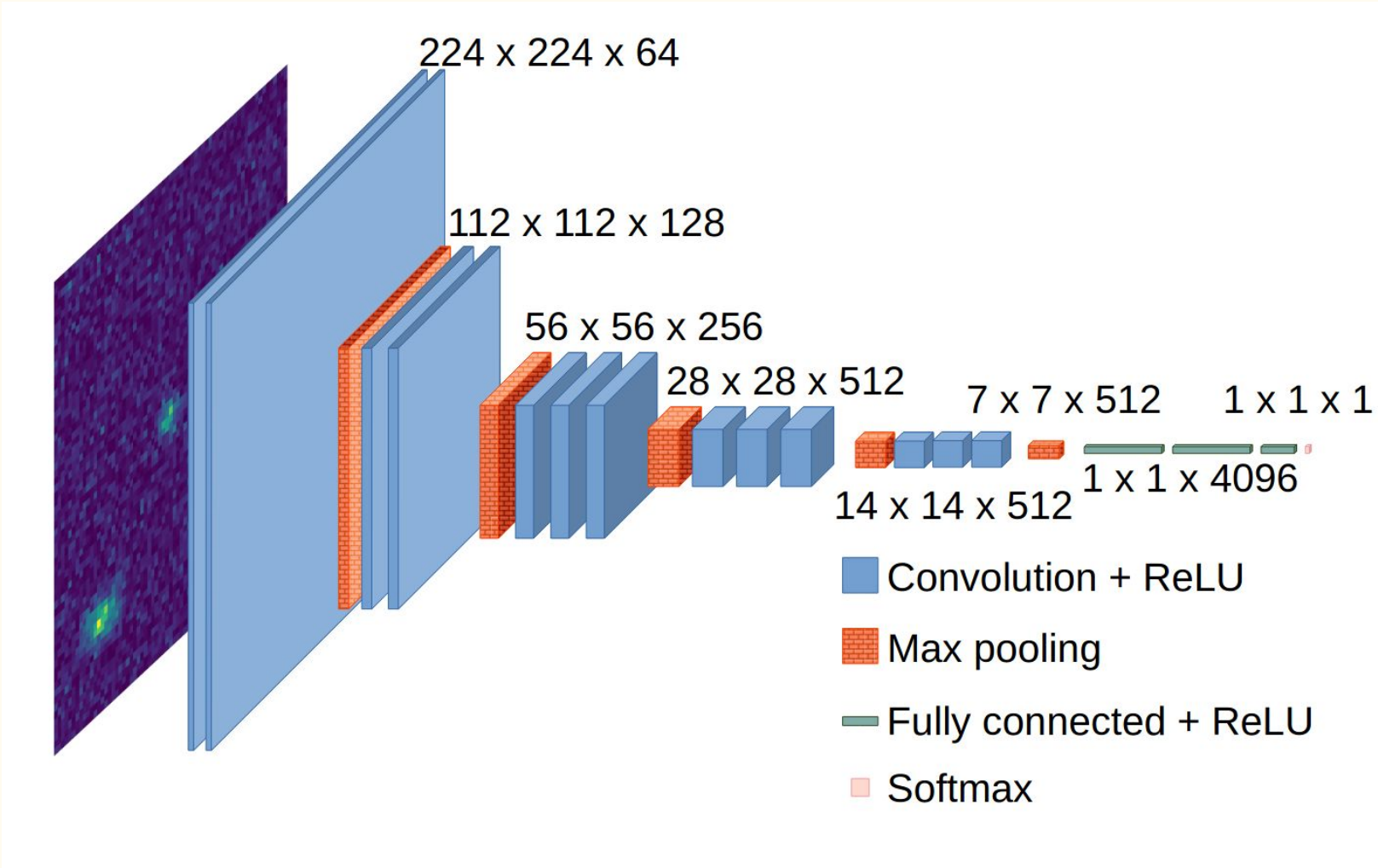



http://towardsdatascience.com

# O que é um filtro de *pooling*?



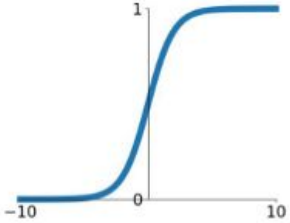

http://towardsdatascience.com


224 x 224 x 64
112 x 112 x 128
56 x 56 x 256
28 x 28 x 512
14 x 14 x 512
7 x 7 x 512
1 x 1 x 4096
1 x 1 x 1
Convolution + ReLU
Max pooling
Fully connected + ReLU
Softmax

# Activation Functions

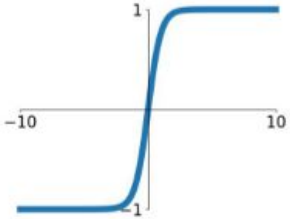
**Sigmoid**

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



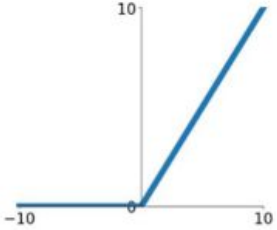

**tanh**

$$\tanh(x)$$



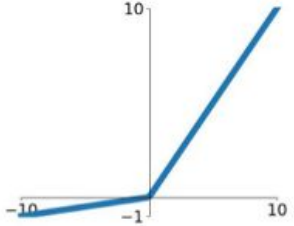

**ReLU**

$$\max(0, x)$$




**Leaky ReLU**

$$\max(0.1x, x)$$

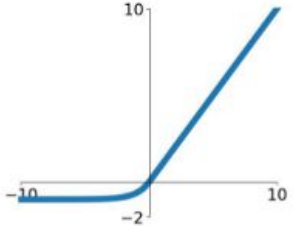



**Maxout**

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

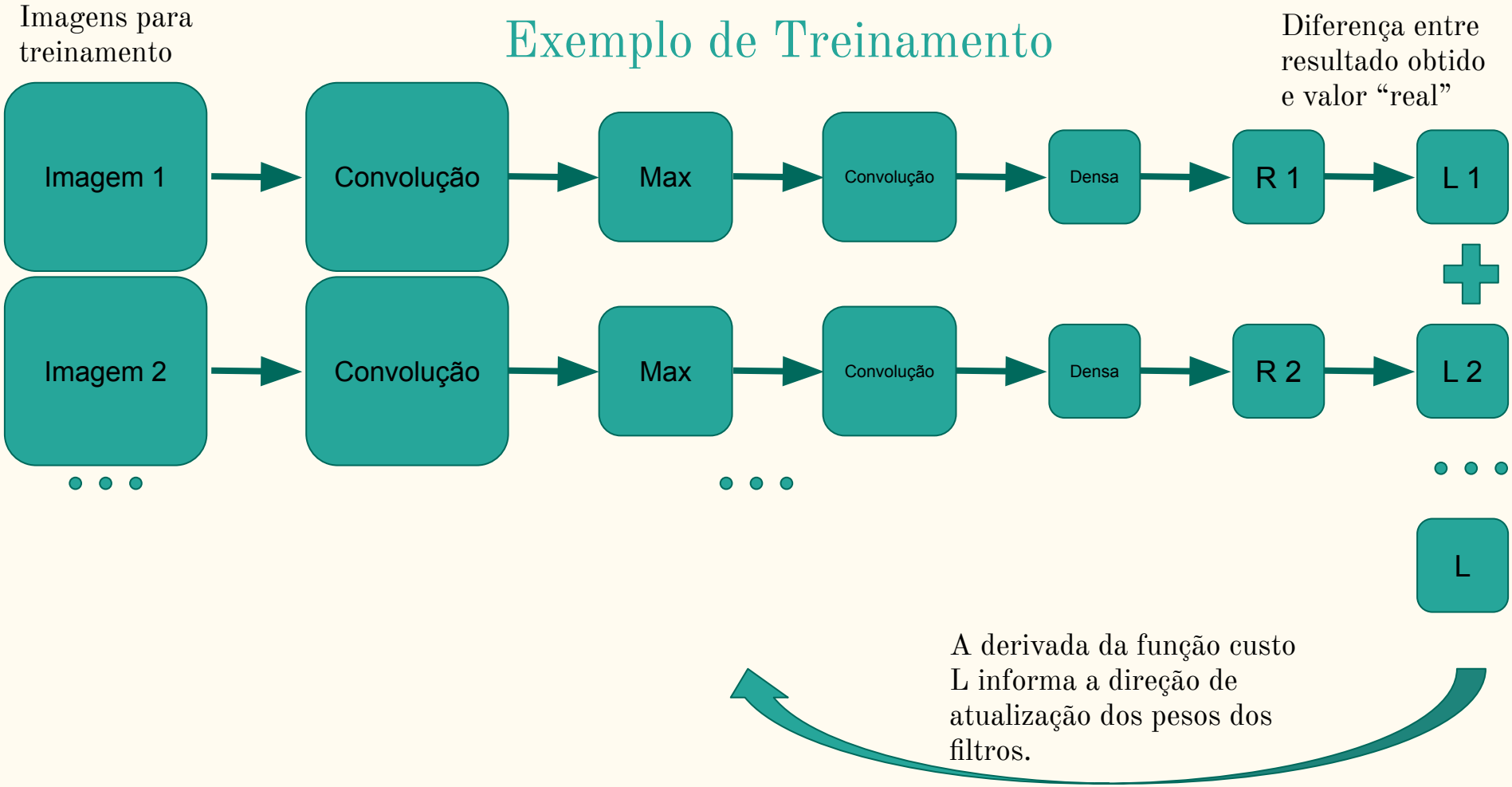
**ELU**

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$




http://towardsdatascience.com

# Exemplo de Treinamento



## Dados préavaliados

### Treino

Dados usados diretamente no processo de treinamento: é neles que se aplica o processo de otimização da rede.

### Validação

Dados usados para avaliar os vários parâmetros de treino.

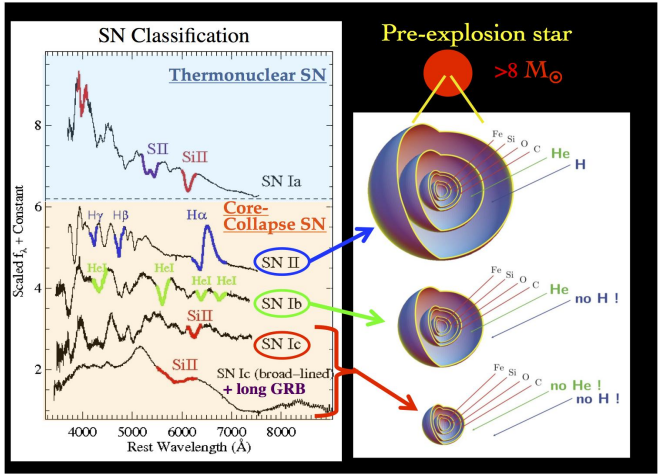
### Teste

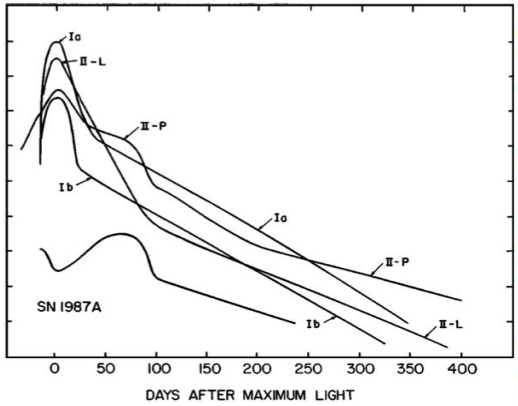
Dados nos quais se faz a avaliação **FINAL** da qualidade da rede treinada. Não podem ser reutilizados até se conseguir a qualidade desejada, senão se tornam automaticamente "validação".

# Dados alvo

Dados sem classificação ou medida anterior, nos quais pretendemos usar nosso modelo treinado para obter novos dados.

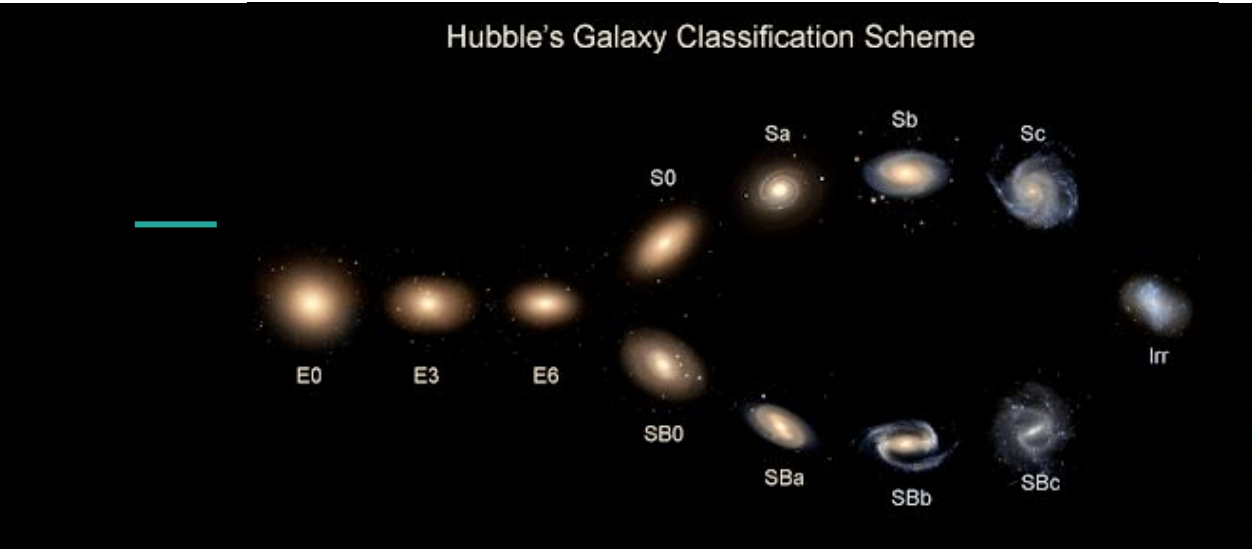
# Aplicações

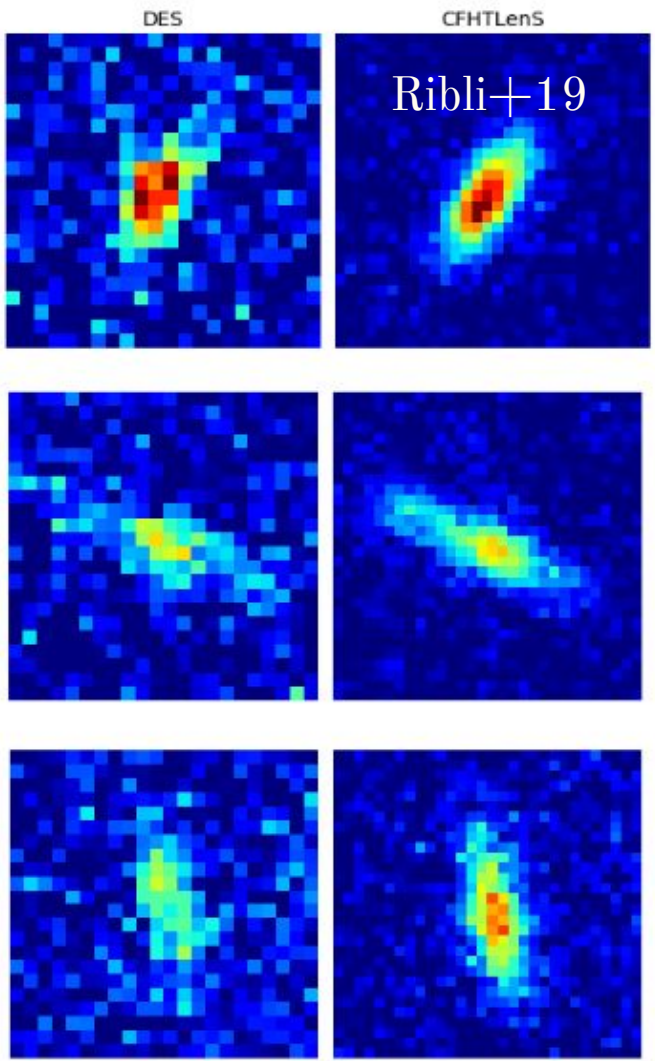


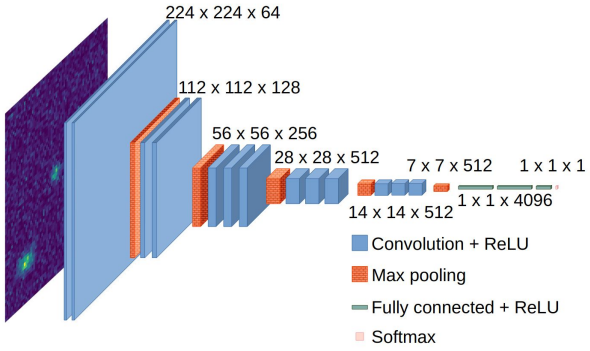




*ure 3* Schematic light curves for SNe of Types Ia, Ib, II-L, II-P, and SN 1987A. The curve SNe Ib includes SNe Ic as well, and represents an average. For SNe II-L, SNe 1979C and 80K are used, but these might be unusually luminous.

*Figure Credit*: Wheeler, J. C., & Harkness, R. P. 1990, RPPh, 53, 1467




224 x 224 x 64
112 x 112 x 128
56 x 56 x 256
28 x 28 x 512
7 x 7 x 512
1 x 1 x 1
1 x 1 x 4096
14 x 14 x 512
Convolution + ReLU
Max pooling
Fully connected + ReLU
Softmax

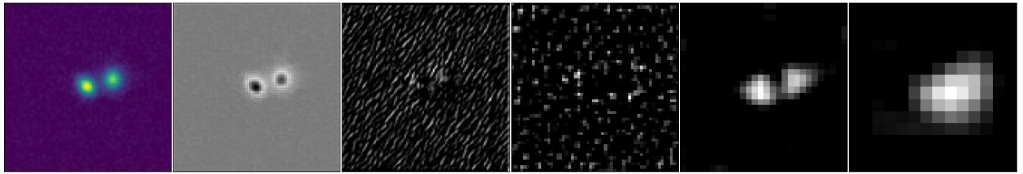
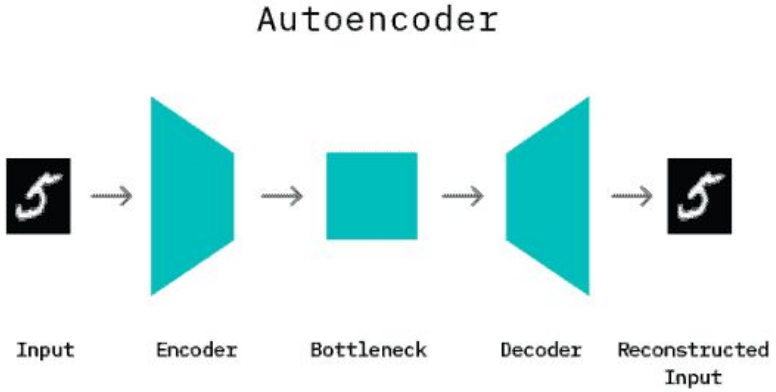


Image (65x65)
Block 1 (224x224)
Block 2 (112x112)
Block 3 (56x56)
Block 4 (28x28)
Block 5 (14x14)




$$p(\boldsymbol{\theta}|\mathbf{d}_o, \mathcal{M}) = \frac{p(\mathbf{d}_o|\boldsymbol{\theta}, \mathcal{M})\ p(\boldsymbol{\theta}|\mathcal{M})}{p(\mathbf{d}_o|\mathcal{M})}$$

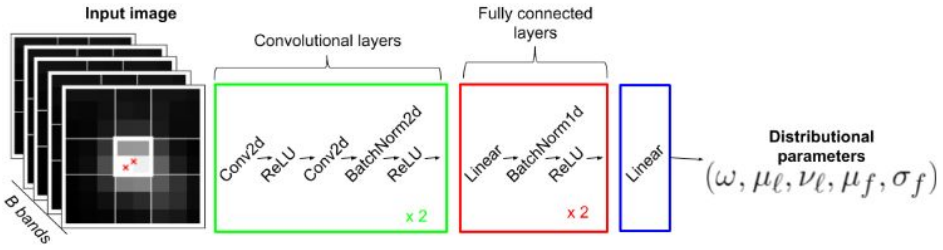
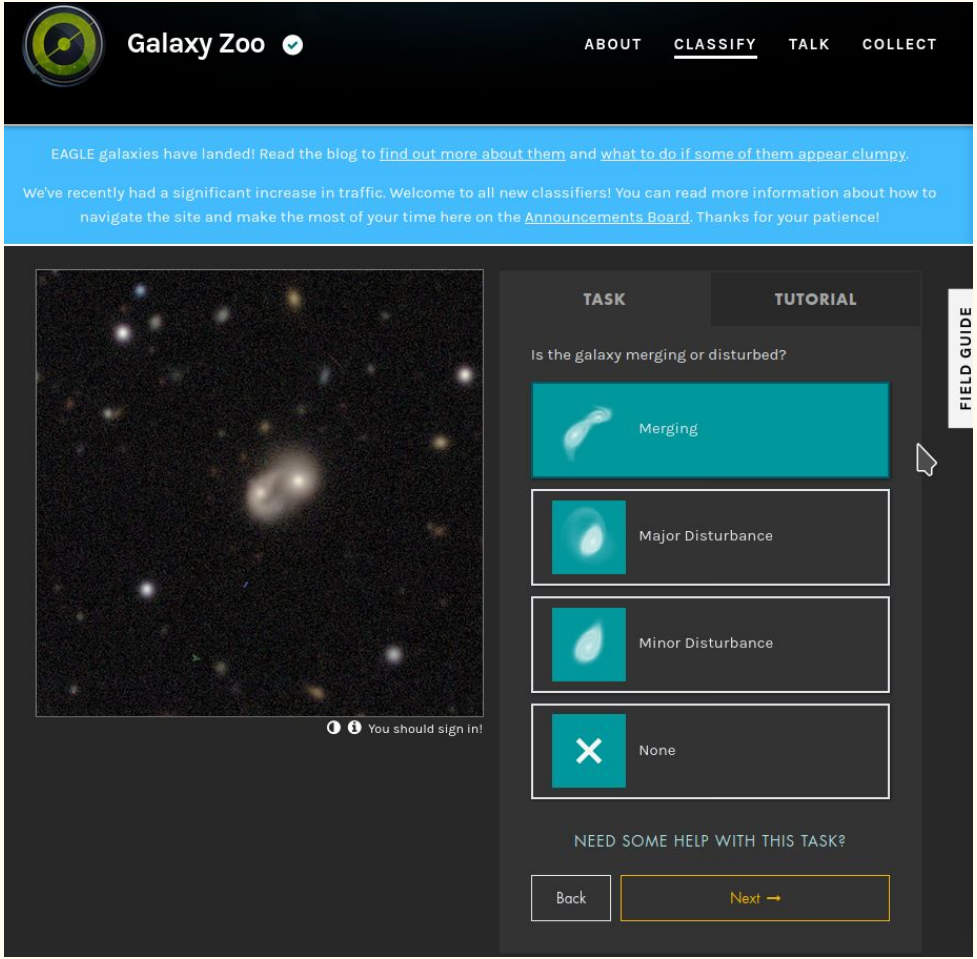



Figure 3: The neural network architecture. For cataloging M2, the input image is an $8 \times 8$ padded tile, and the network returns distributional parameters for latent variables contained in the center $2 \times 2$ tile.

# Dificuldades

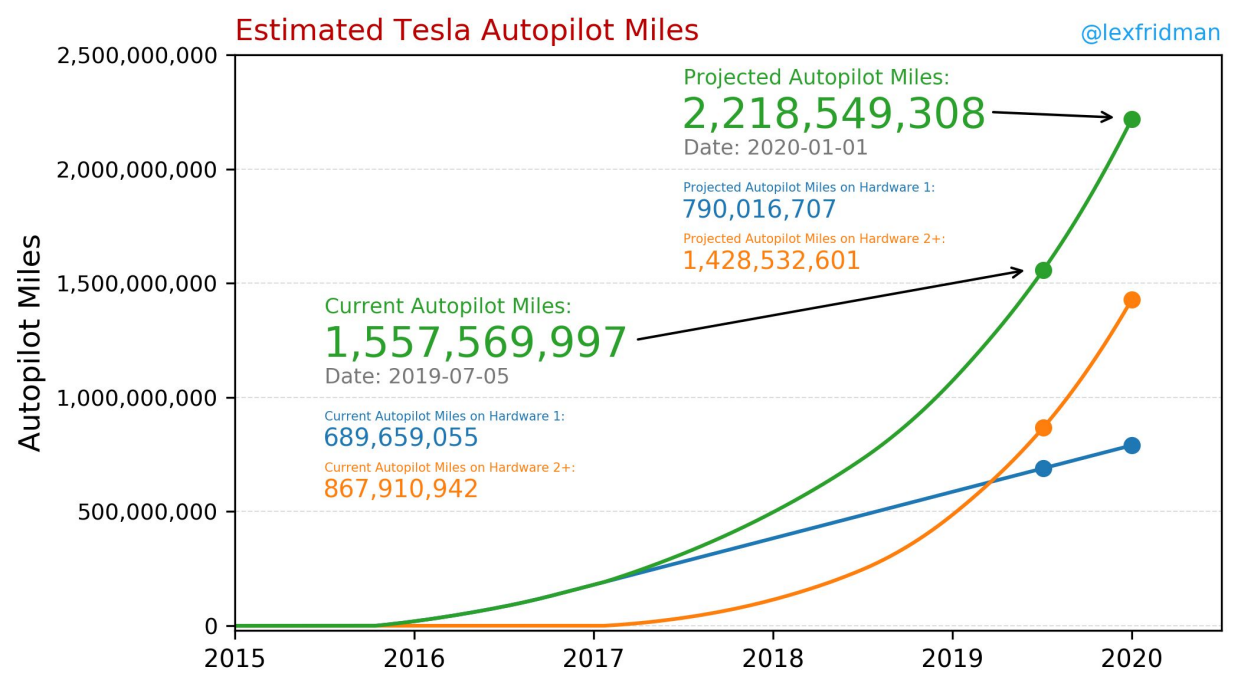
# Dificuldades

1. Dados - pré classificação



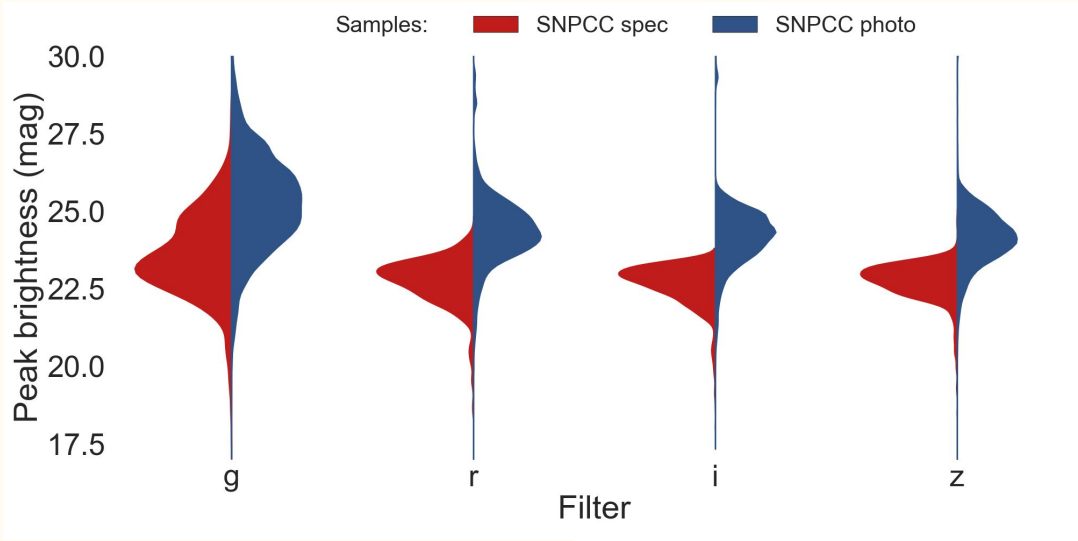
# Dificuldades

1. Dados - pré classificação
2. Dados - volume

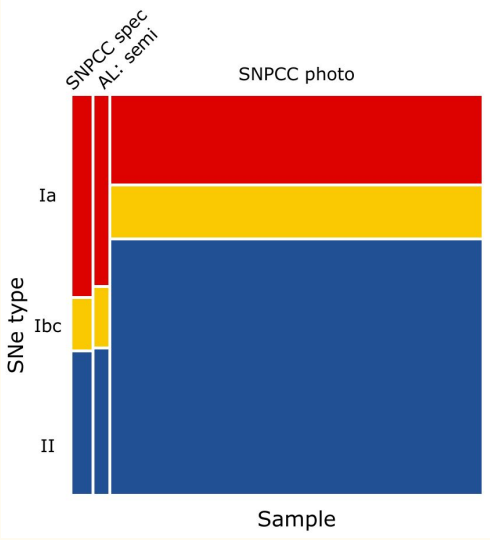


# Dificuldades

1. Dados - pré classificação
2. Dados - volume
3. Dados - representatividade

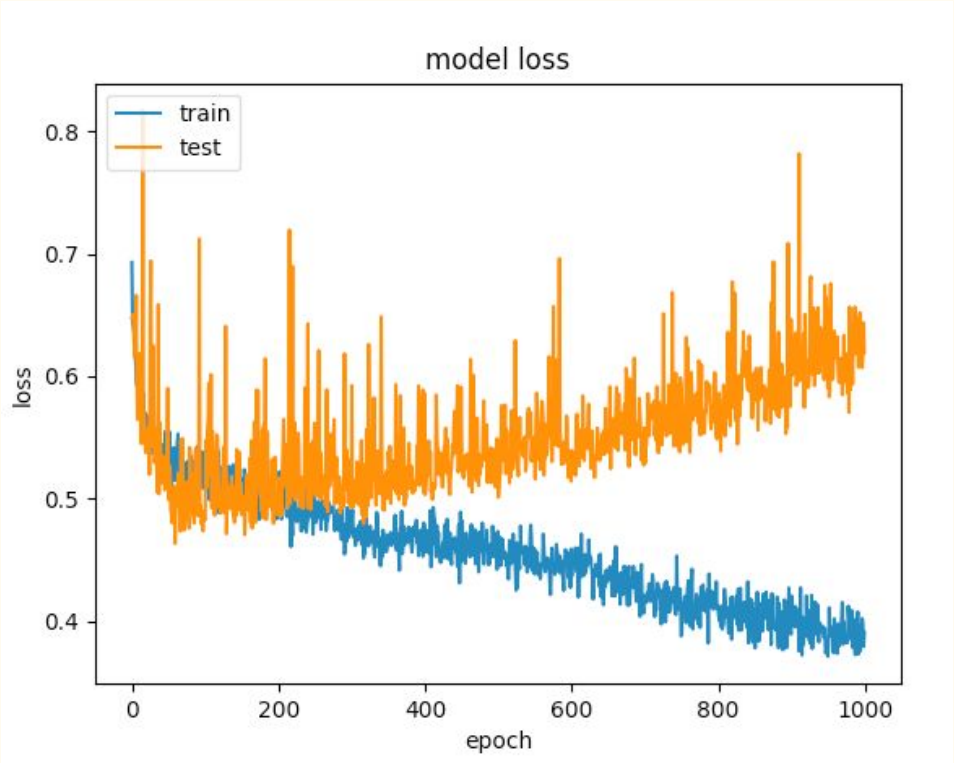



Ishida+19



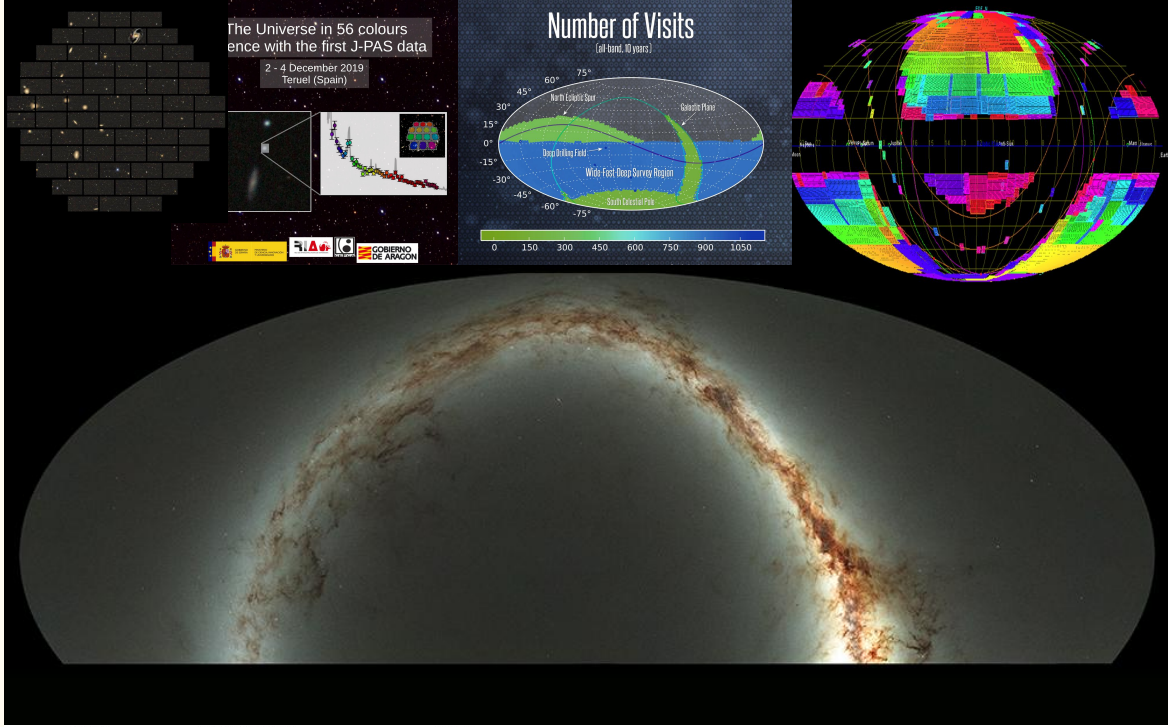
# Dificuldades

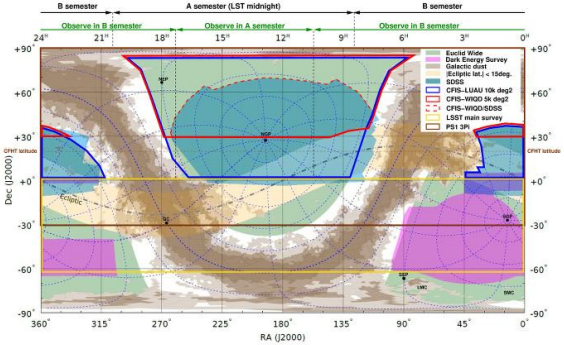
1. Dados - pré classificação
2. Dados - volume
3. Dados - representatividade
4. Overfitting



# Dificuldades

1. Dados - pré classificação
2. Dados - volume
3. Dados - representatividade
4. Overfitting
5. Reusabilidade



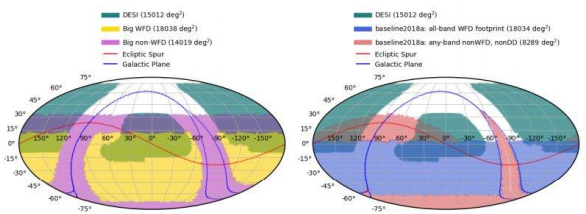




Figure 1: *Left:* Our proposed Big Sky footprint: yellow fields denote our recommended expanded WFD footprint while the purple fields represent the mini-surveys in the extended footprint. *Right:* Footprint from `baseline2018a` for WFD (blue) and all the mini-surveys aside from the DDFs (coral red). Both plots show overlap the DESI footprint (aqua green), demonstrating that our Big Sky footprint significantly increases the overlap with DESI (5912 deg$^2$ for WFD and 4538 deg$^2$ for non-WFD) vs. `baseline2018a` (3739 deg$^2$ for WFD and 2233 deg$^2$ for non-WFD).

# Dificuldades

1. Dados - pré classificação
2. Dados - volume
3. Dados - representatividade
4. Overfitting
5. Reusabilidade
6. Determinação de erros

# Dificuldades

1. Dados - pré classificação
2. Dados - volume
3. Dados - representatividade
4. Overfitting
5. Reusabilidade
6. Determinação de erros
7. Simetrias

# Como usar Deep Learning em imagens Astronômicas

# Frameworks

**Tensorflow, PyTorch**

- Mais genérico, flexível
- Mais complexidade
- Melhor performance

**Keras**

- Alto nível
- Simples
- Relativamente mais lento

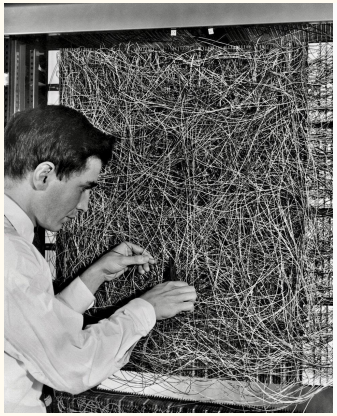
# Extra

# Para saber mais

1. Tutoriais de Deep Learning & Python: sentdex (YouTube)
2. Como funciona: 3Blue1Brown (legendado em BR/PT)
3. Curiosidades, insights: Lex Fridman
4. Keras - https://keras.io/
5. https://towardsdatascience.com/
6. Datasets/Competições: Kaggle

# Histórico

1. Perceptron (Rosenblatt 1957)
2. Neocognitron (Fukushima 1980)
3. Backpropagation, CNNs (LeCun 1986)
4. GPUs (2009)
5. Tensorflow, Pytorch, Keras...



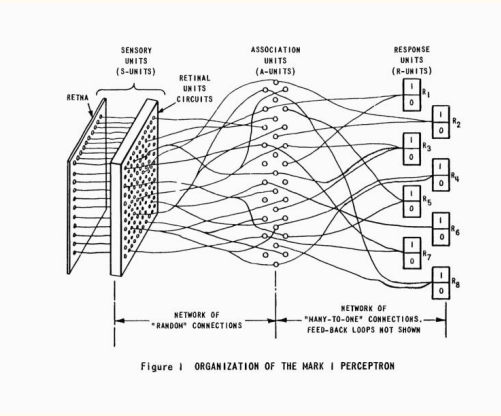


Figure 1 ORGANIZATION OF THE MARK I PERCEPTRON