Conoscere l'Intelligenza Artificiale

Alberto Rota
Francesca Fati

Lezione 2
AI Tasks, Overfitting and Applications

Agenda

Lezione 1

- Introduzione a AI
- Neuroni e Perceptron
- Training
- Importanza dei Dati + Pulizia

→ **Pratica: SPAZIO**

Lezione 2

- Cosa fare con l'AI
- Applicazioni
- Loss Functions
- Iperparametri
- Overfitting e Generalizzazione

→ **Pratica: MELE**

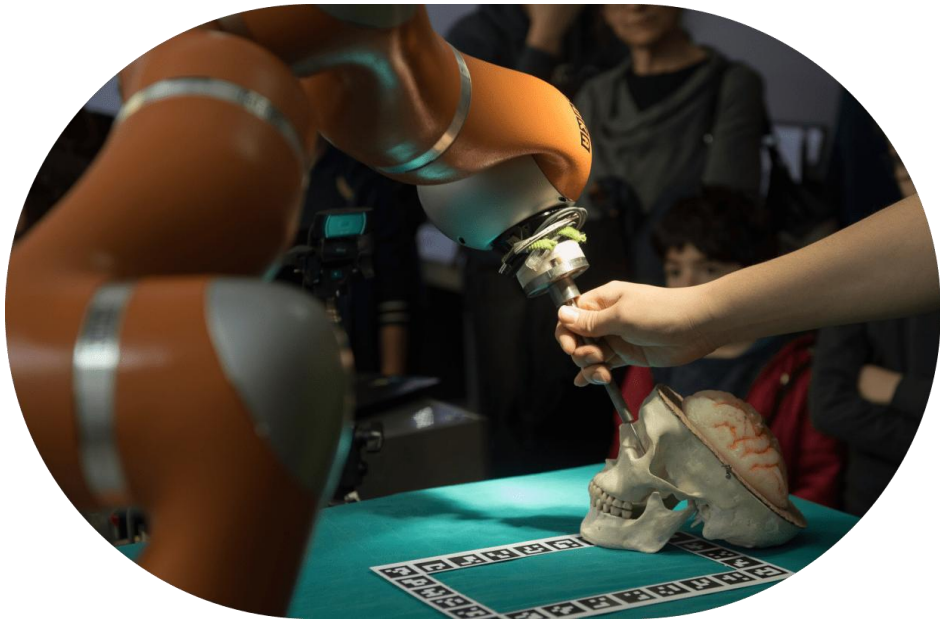
Lezione 3

- ChatGPT
- Privacy / AI Act
- Explainability
- Competition!

→ **TEMA: ?**

Lezione 4

LEONARDO LABS
ROBOTICS
POLITECNICO MILANO 1863 - DEIB



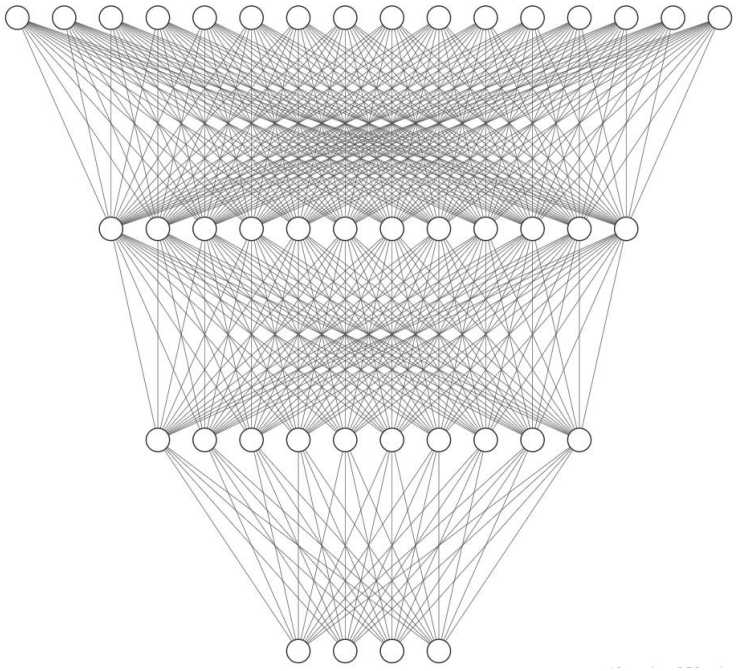
Recap

I Tre componenti principali dell'Intelligenza Artificale

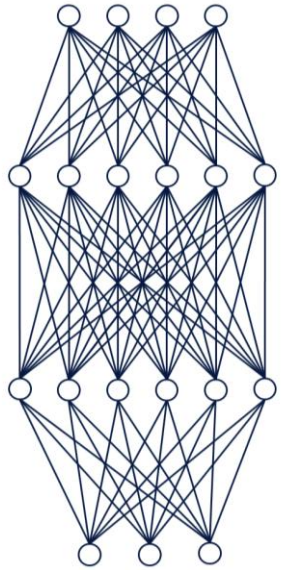
Dataset

Cust City	Cust State	Cust Zip Code	Cust County	Cust Region
Costa Mesa	California	92628	Orange County, California	West
Edinburg	Texas	78539	Hidalgo County, Texas	Southwest
Vancouver	Washington	98668	Clark County, Washington	West
El Paso	Texas	79910	El Paso County, Texas	Southwest
Columbus	Georgia	31908	Muscogee County, Georgia	South
Omaha	Nebraska	68108	Douglas County, Nebraska	Midwest
Mountain View	California	94035	Contra Costa County, Californ	West
Lincoln	Nebraska	68501	Lancaster County, Nebraska	Midwest
Tulsa	Oklahoma	74107	Osage County, Oklahoma	Southwest
Gastonia	North Carolina	28052	Gaston County, North Carolin	South
Newport News	Virginia	23501	Newport News City, Virginia	East
Dayton	Ohio	45401	Montgomery County, Ohio	Midwest
South Gate	California	90280	Los Angeles County, Californi	West
Birmingham	Alabama	35203	Jefferson County, Alabama	South
Sacramento	California	95813	Sacramento County, Californi	West
Dallas	Texas	75260	Dallas County, Texas	Southwest
Antioch	California	94509	Contra Costa County, Californ	West
Lubbock	Texas	79402	Lubbock County, Texas	Southwest
Bellflower	California	90706	Los Angeles County, Californi	West
St. Paul	Minnesota	55109	Ramsey County, Minnesota	Midwest
Mesquite	Texas	75149	Dallas County, Texas	Southwest
Rio Rancho	New Mexico	87124	Bernalillo County, New Mexic	Southwest
Edinburg	Texas	78539	Hidalgo County, Texas	Southwest
Camden	New Jersey	08101	Camden County, New Jersey	East
Sioux Falls	South Dakota	57104	Lincoln County, South Dakota	Midwest
Lexington-Fayette	Kentucky	40511	Fayette County, Kentucky	South
Louisville	Kentucky	40231	Jefferson County, Kentucky	South
Fall River	Massachusetts	02720	Bristol County, Massachusett	East
Palmdale	California	93550	Los Angeles County, Californi	West
Sacramento	California	95813	Sacramento County, Californi	West
Sioux Falls	South Dakota	57104	Lincoln County, South Dakota	Midwest
Austin	Texas	78710	Travis County, Texas	Southwest
Quincy	Massachusetts	02169	Norfolk County, Massachuset	East
Roswell	Georgia	30075	Fulton County, Georgia	South
Oakland	California	94612	Alameda County, California	West
Bellflower	California	90706	Los Angeles County, Californi	West
Phoenix	Arizona	85026	Maricopa County, Arizona	Southwest

Modello

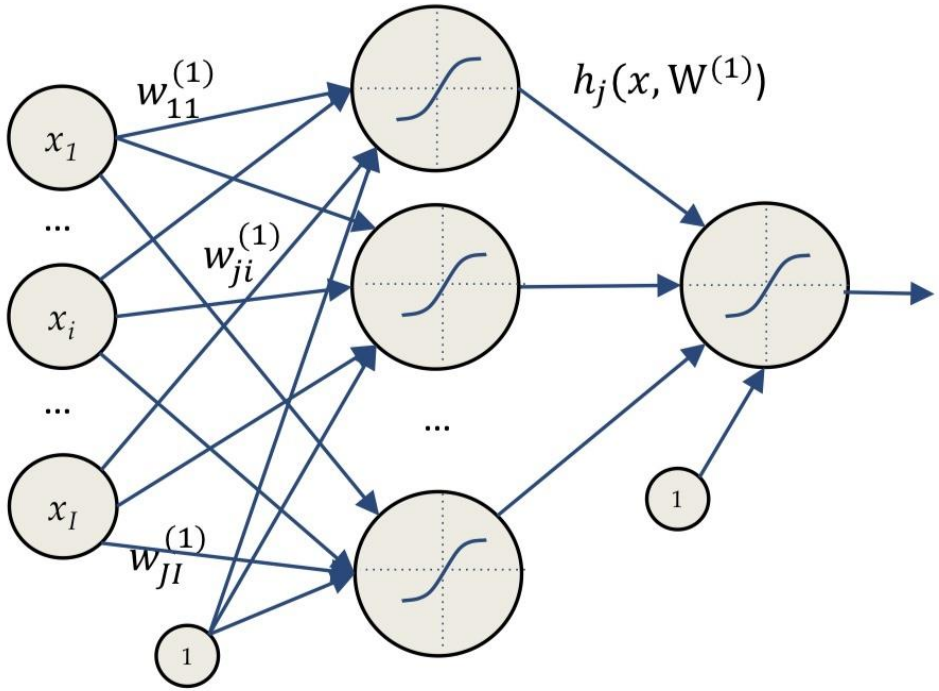


Ottimizzazione



Remarks

Gradient Descent Example

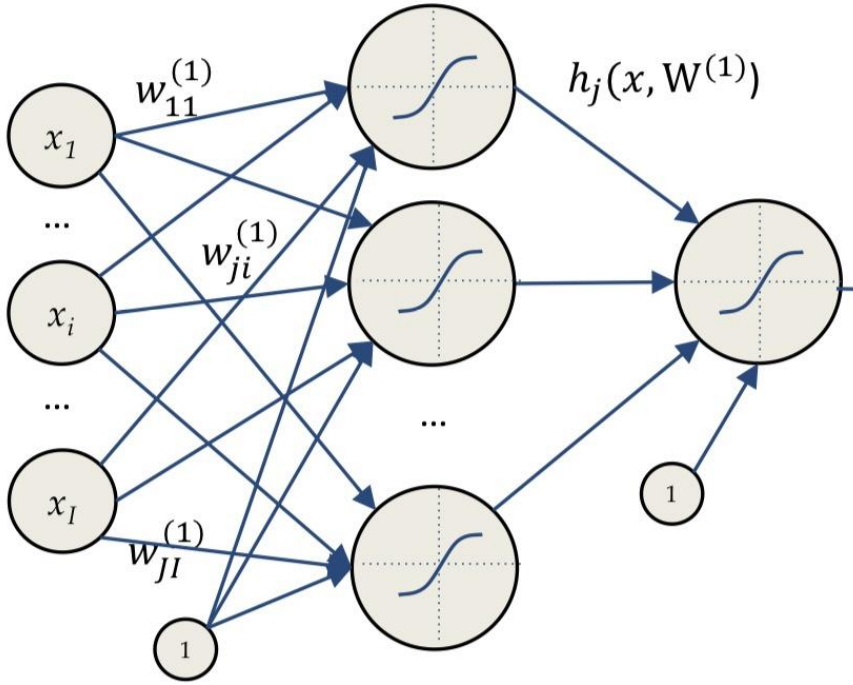


$$g_1(x_n|w) = g_1\left(\sum_{j=0}^{J} w_{1j}^{(2)} \cdot h_j\left(\sum_{i=0}^{I} w_{ji}^{(1)} \cdot x_{i,n}\right)\right)$$

$$E(w) = \sum_{n=1}^{N} (t_n - g_1(x_n, w))^2$$

aken from Prof. Matteo Matteucci

Gradient Descent Example



$$g_1(x_n|w) = g_1\left(\sum_{j=0}^{J} w_{1j}^{(2)} \cdot h_j\left(\sum_{i=0}^{I} w_{ji}^{(1)} \cdot x_{i,n}\right)\right)$$

$$E(w) = \sum_{n=1}^{N} (t_n - g_1(x_n, w))^2$$

$$\frac{\partial E(w_{ji}^{(1)})}{\partial w_{ji}^{(1)}} = -2 \sum_{n}^{N} \left(t_n - g_1(x_n, w)\right) g_1'(x_n, w) w_{1j}^{(2)} h_j'\left(\sum_{i=0}^{I} w_{ji}^{(1)} \cdot x_{i,n}\right) x_i$$

aken from Prof. Matteo Matteucci

$$g_1(x_n|w) = g_1\left(\sum_{j=0}^{J} w_{1j}^{(2)} \cdot h_j\left(\sum_{i=0}^{I} w_{ji}^{(1)} \cdot x_{i,n}\right)\right)$$

$$E(w) = \sum_{n=1}^{N} (t_n - g_1(x_n, w))^2$$

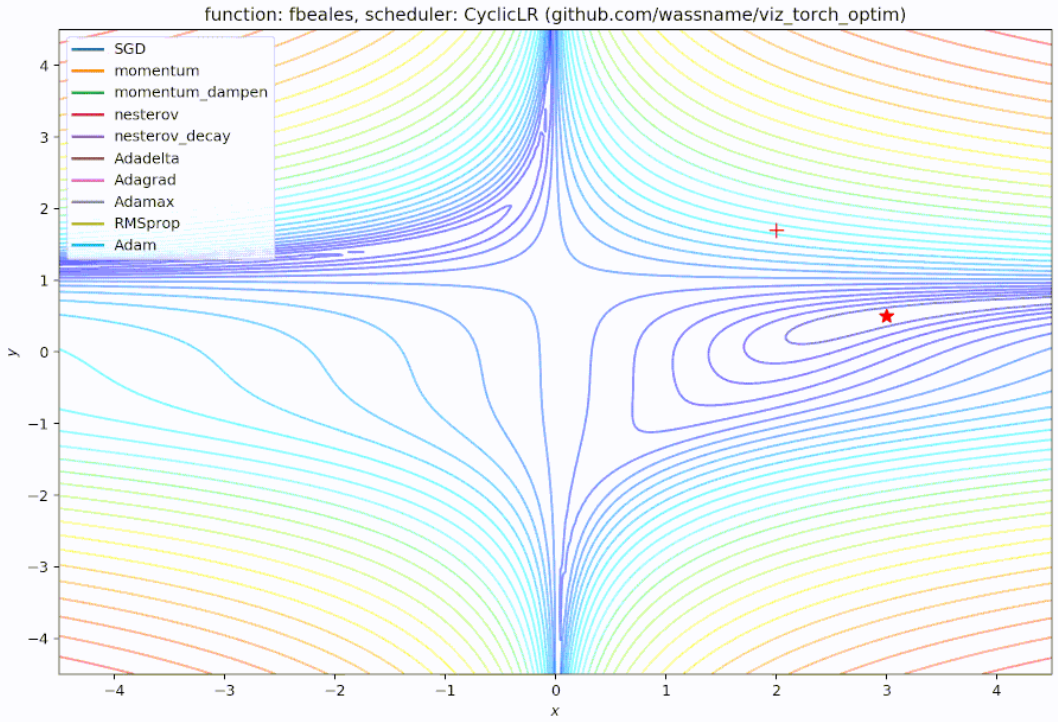
$$\frac{\partial E(w)}{\partial w_{3,5}^{(1)}} = \frac{\partial \sum_{n=1}^{N} (t_n - g_1(x_n, w))^2}{\partial w_{3,5}^{(1)}} = \sum_{n=1}^{N} \frac{\partial (t_n - g_1(x_n, w))^2}{\partial w_{3,5}^{(1)}} = -2 \sum_{n}^{N} \left(t_n - g_1(x_n, w)\right) \frac{\partial g_1(x_n, w)}{\partial w_{3,5}^{(1)}}$$

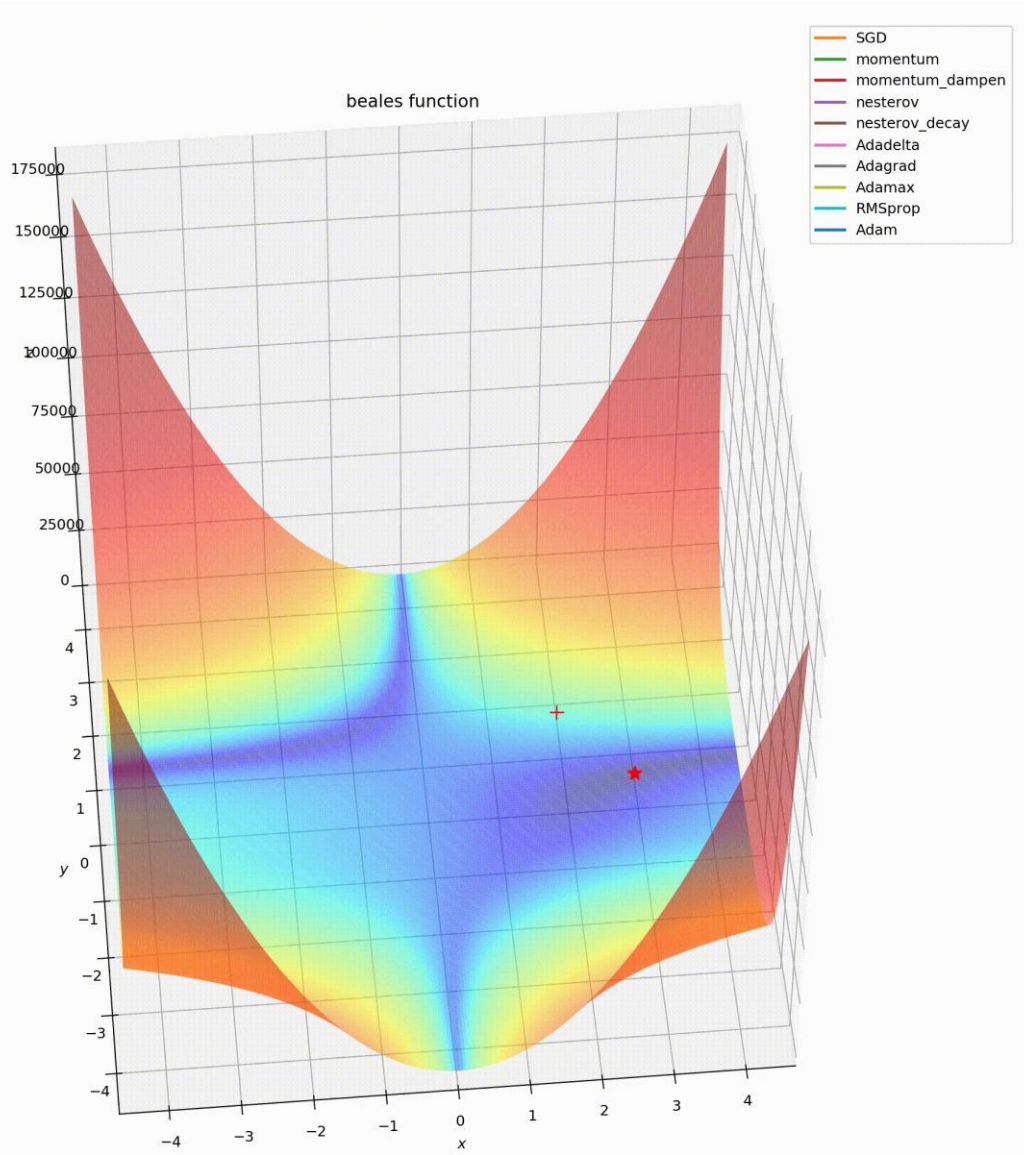
$$\frac{\partial g_1(x_n, w)}{\partial w_{3,5}^{(1)}} = \frac{\partial g_1\left(\sum_{j=0}^{J} w_{1j}^{(2)} \cdot h_j(.)\right)}{\partial w_{3,5}^{(1)}} = g_1'(x_n, w) \cdot \frac{\partial \sum_{j=0}^{J} w_{1j}^{(2)} \cdot h_j(.)}{\partial w_{3,5}^{(1)}} = g_1'(x_n, w) \cdot w_{1,3}^{(2)} \cdot \frac{\partial h_3\left(\sum_{i=0}^{I} w_{3i}^{(1)} \cdot x_{i,n}\right)}{\partial w_{3,5}^{(1)}}$$

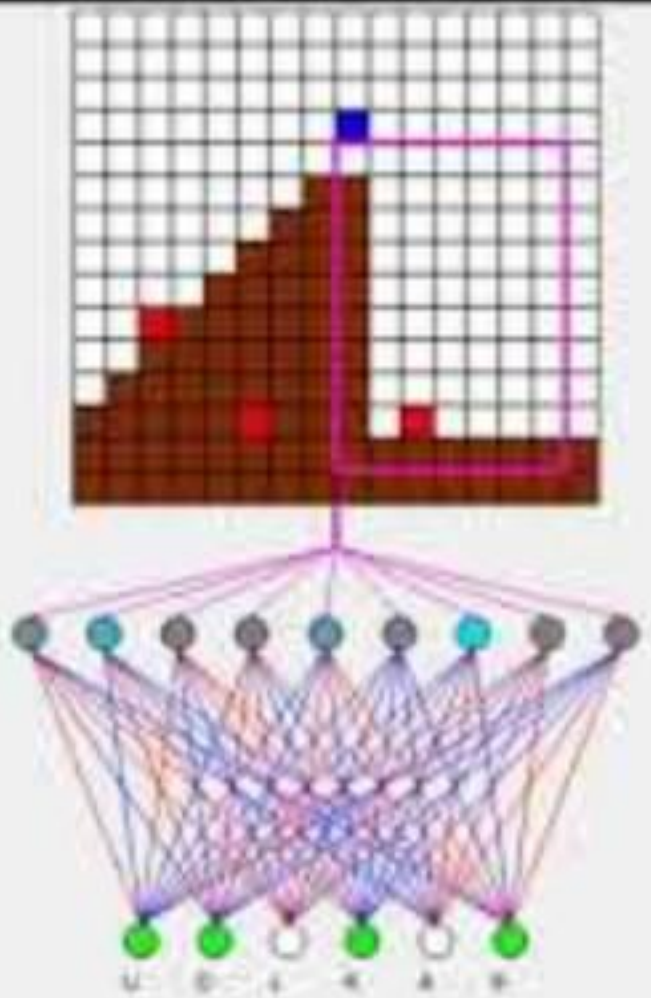
$$\frac{\partial h_3\left(\sum_{i=0}^{I} w_{3i}^{(1)} \cdot x_{i,n}\right)}{\partial w_{3,5}^{(1)}} = h_3'\left(\sum_{i=0}^{I} w_{3,i}^{(1)} \cdot x_{i,n}\right) \frac{\partial \sum_{i=0}^{I} w_{3,i}^{(1)} \cdot x_{i,n}}{\partial w_{3,5}^{(1)}} = h_3'\left(\sum_{i=0}^{I} w_{3,i}^{(1)} \cdot x_{i,n}\right) x_{5,n}$$

$$\frac{\partial E(w)}{\partial w_{3,5}^{(1)}} = -2 \sum_{n}^{N} \left(t_n - g_1(x_n, w)\right) g_1'(x_n, w) w_{1,3}^{(2)} h_3'\left(\sum_{i=0}^{I} w_{3,i}^{(1)} \cdot x_{i,n}\right) x_{5,n}$$

aken from Prof. Matteo Matteucci


function: fbeales, scheduler: CyclicLR (github.com/wassname/viz_torch_optim)
SGD
momentum
momentum_dampen
nesterov
nesterov_decay
Adadelta
Adagrad
Adamax
RMSprop
Adam
x
y


beales function
SGD
momentum
momentum_dampen
nesterov
nesterov_decay
Adadelta
Adagrad
Adamax
RMSprop
Adam
x
y



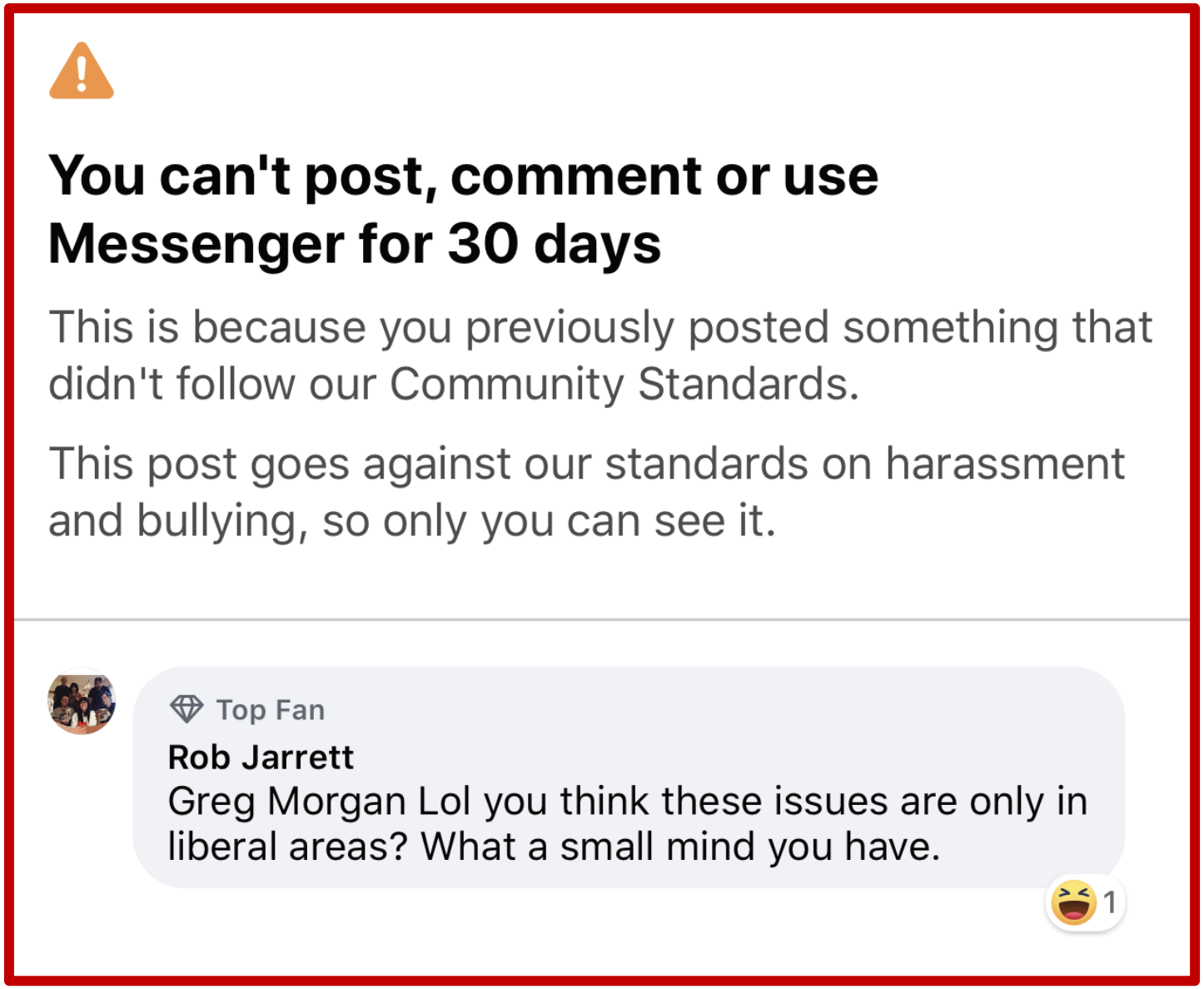
Generation:	[illegible]	Offspring:	10, 10
Individual:	Replay	Lifespan:	Infinite
Best Fitness:	0	Mutation:	Static 5.0%
Max Distance:	2710	Crossover:	Roulette
Num Inputs:	10	SBX Eta:	100.0
Trainable Params:	703	Layers:	[10, 5, 6]

Applicazioni

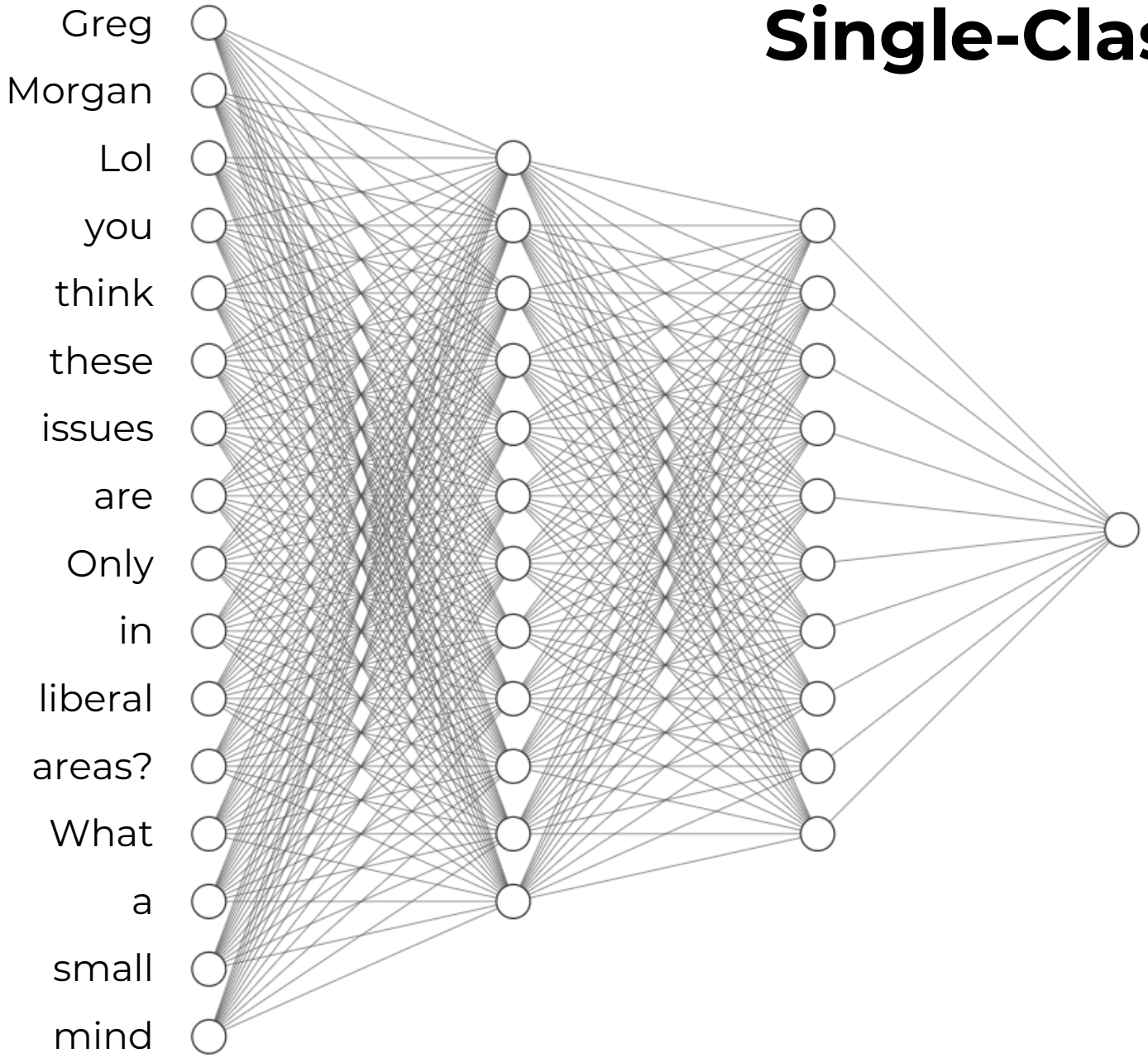
Classification

A quale “gruppo” / “classe” appartiene il mio dato?

- Sano / Malato
- Good / Bad

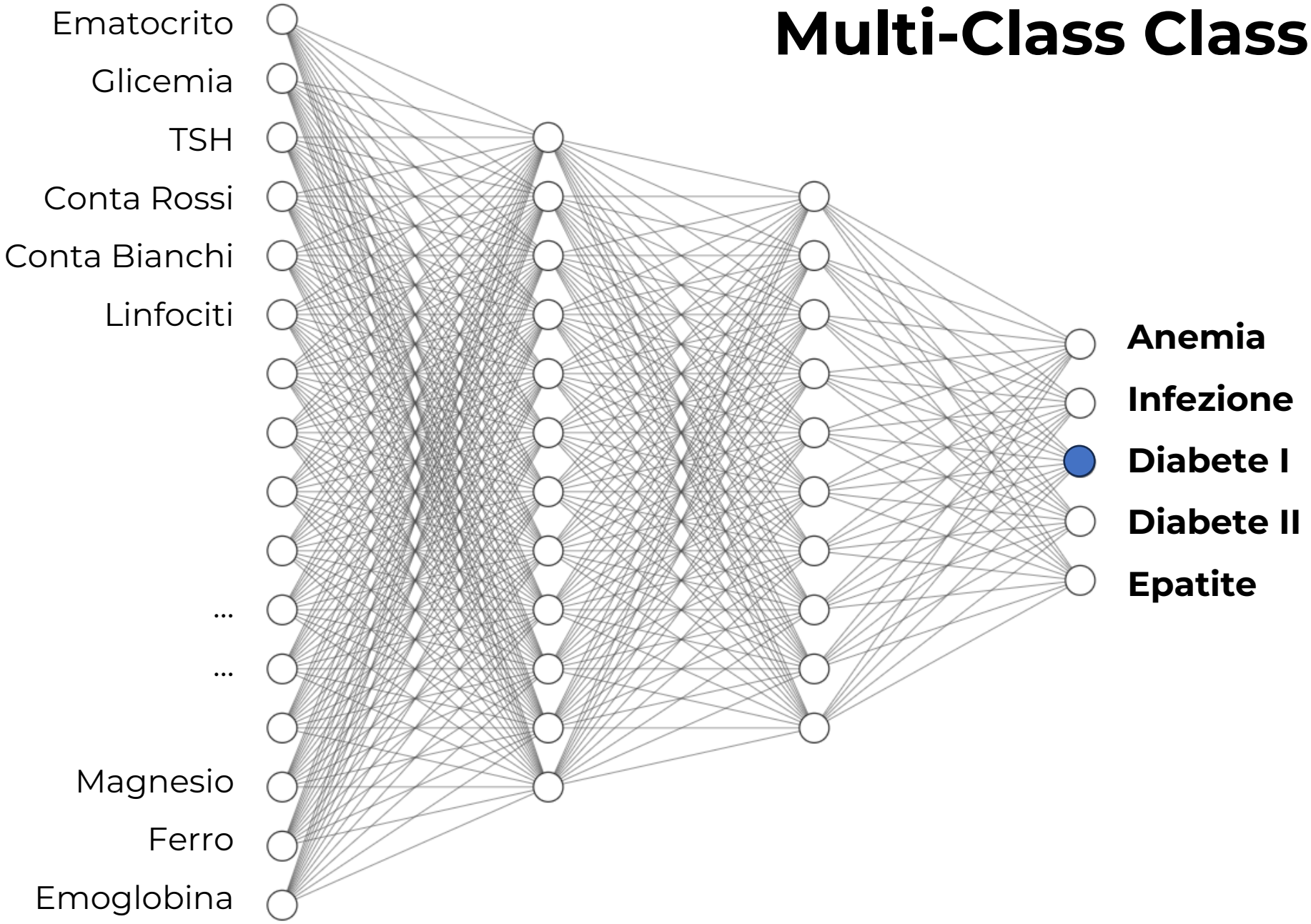


Single-Class Classification

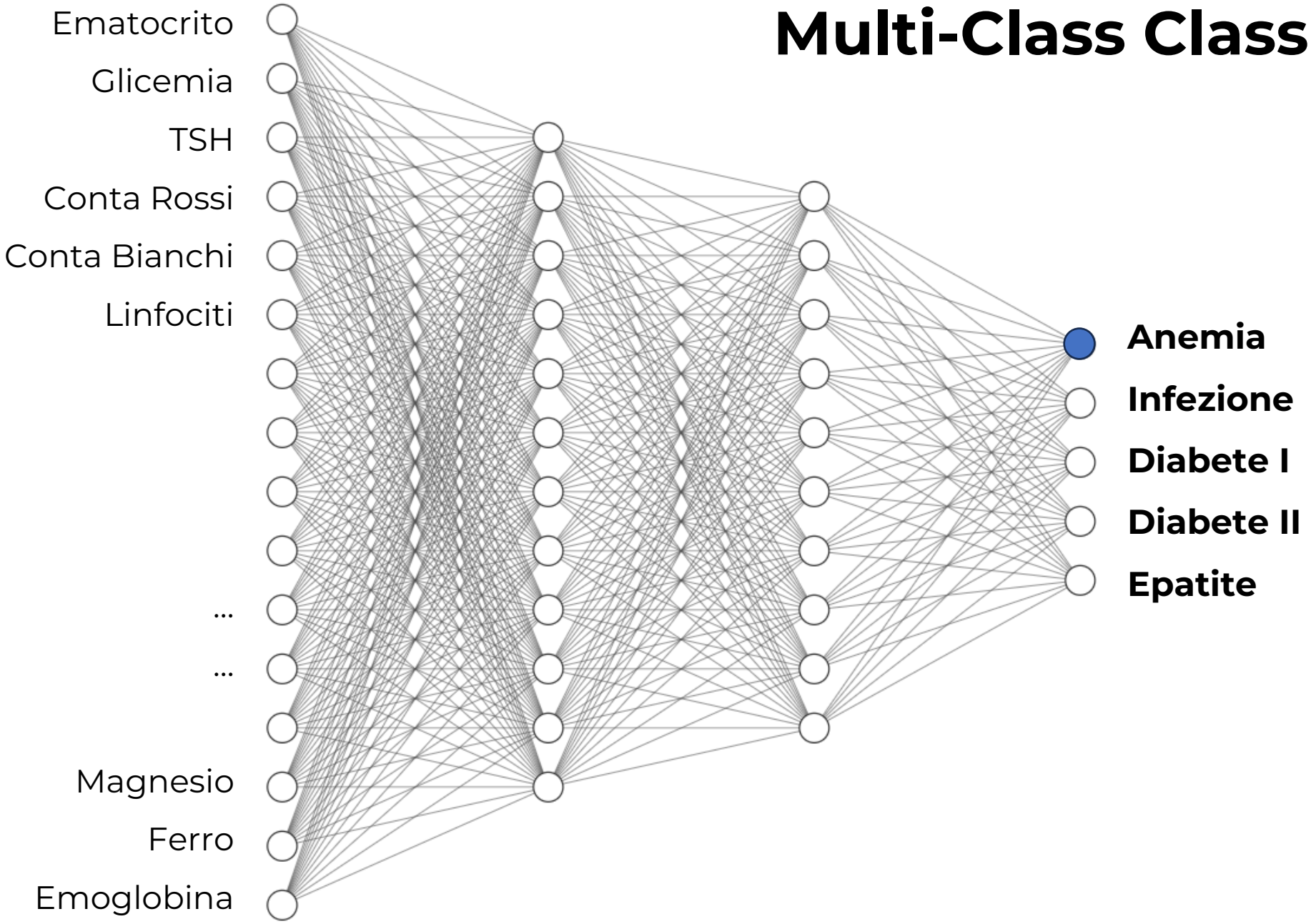


0 = Appropriate
1 = Inappropriate

Multi-Class Classification

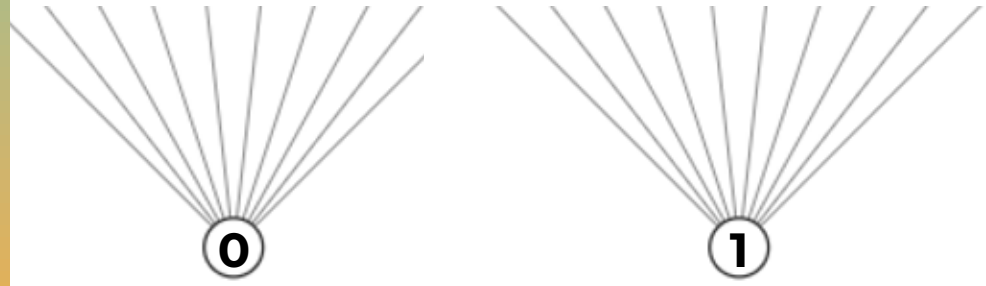


Multi-Class Classification



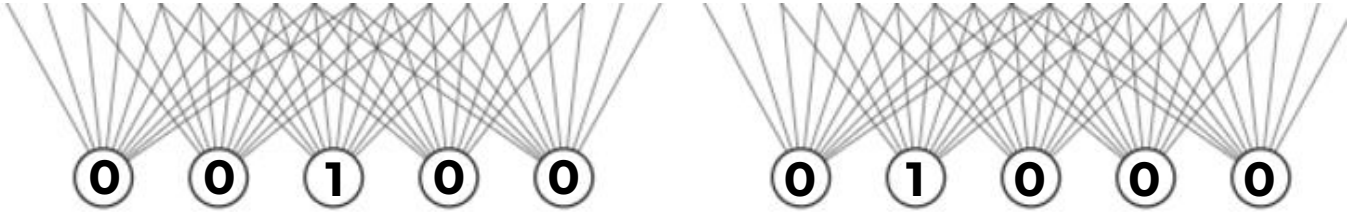
Errore / Loss

Single-Class Classification



$$E = y - \hat{y}$$

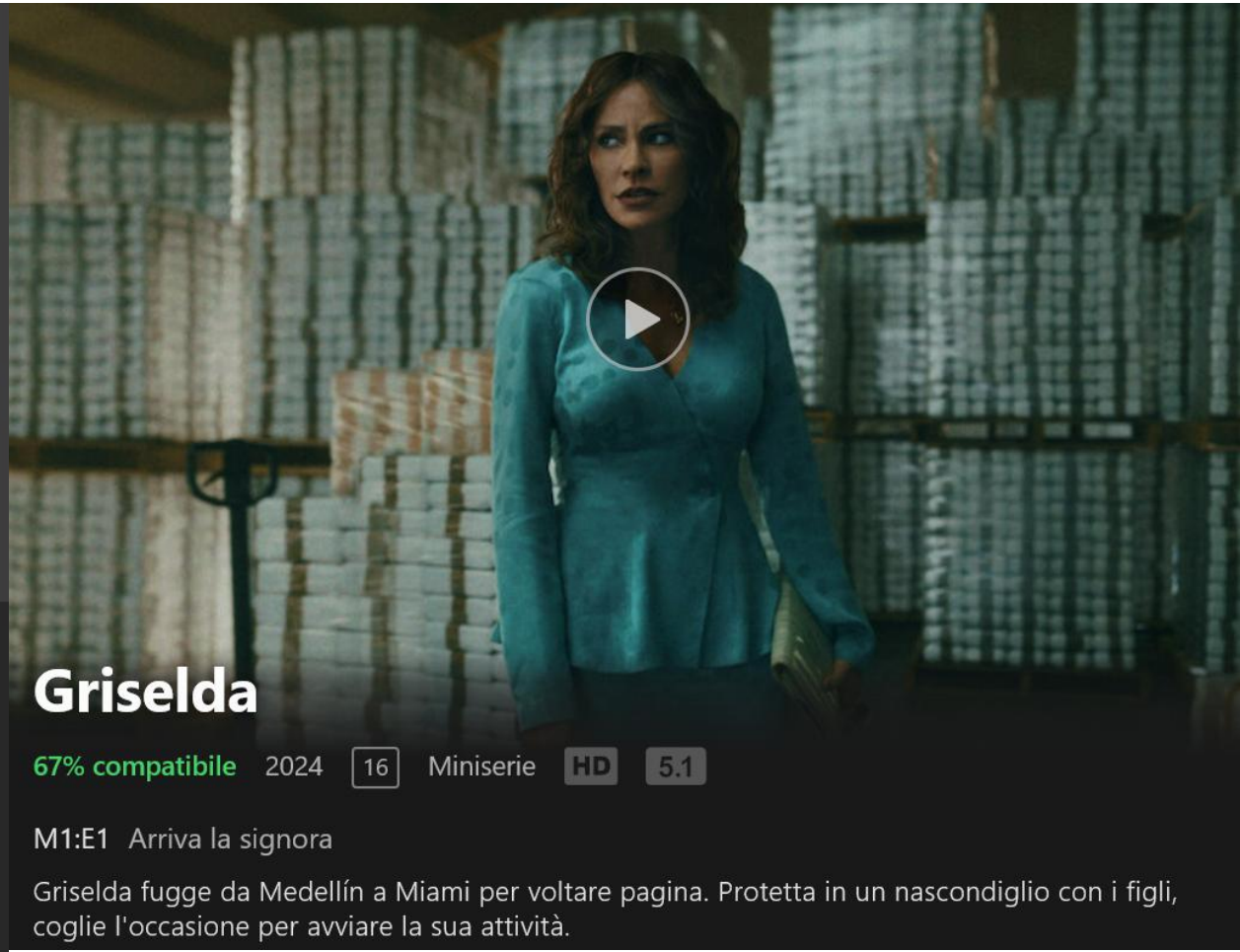
Multi-Class Classification

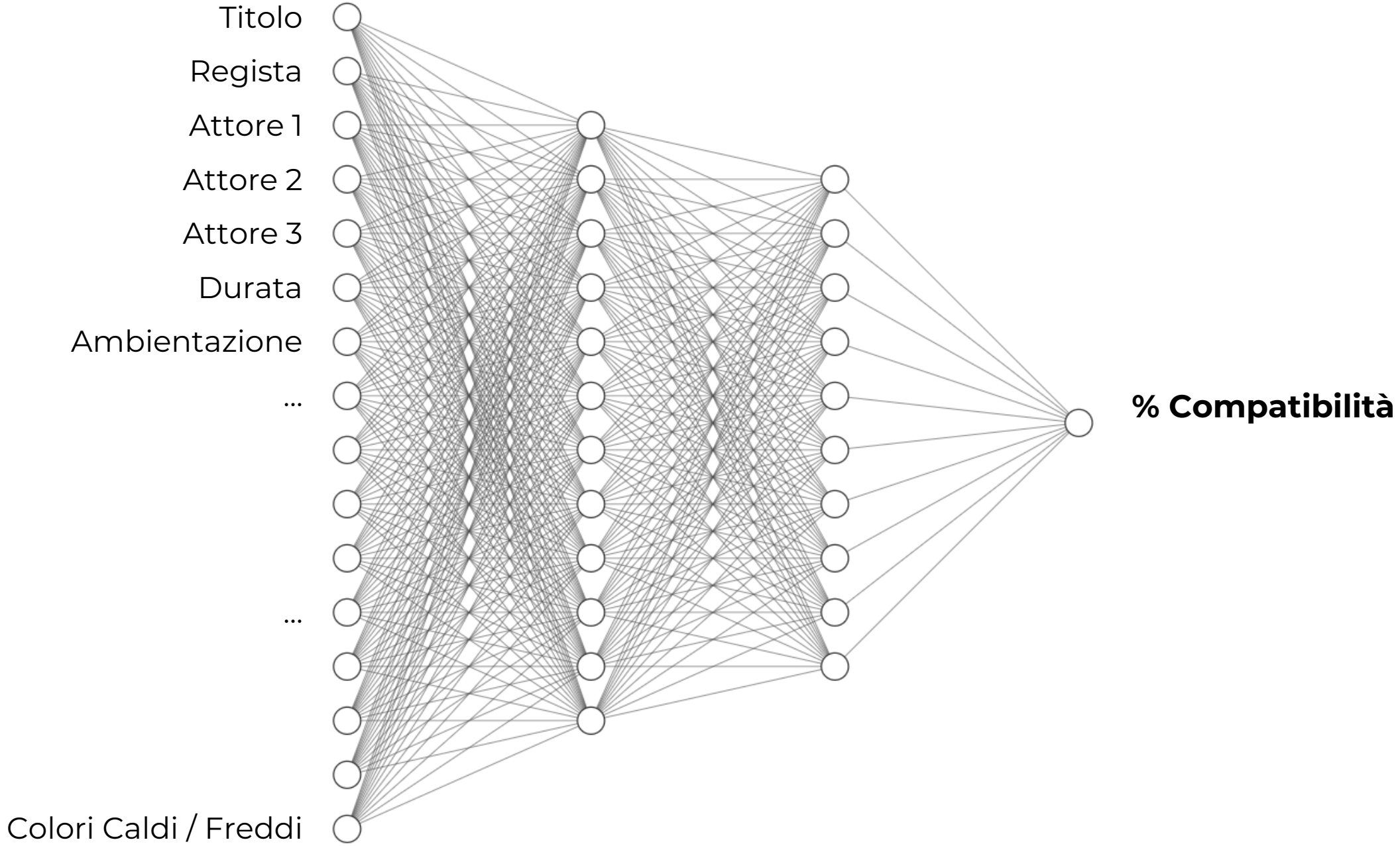


$$CE = -\sum_{i=1}^{i=N} y_i \cdot log(\widehat{y_i})$$

Regression

Quale valore (continuo) è associato al mio dato?




Titolo
Regista
Attore 1
Attore 2
Attore 3
Durata
Ambientazione
...
...
Colori Caldi / Freddi
% Compatibilità

3Bmeteo

lunedì 11 MARZO				
	6°C	5°C	14°C	10°C
Percepita	2°C	1°C	14°C	9°C
Vento	4 Km/H SO	13 Km/H ONO	14 Km/H O	13 Km/H O
Zero Termico	1584	1370	1449	1637
Precipitazioni	-	-	-	-
Umidità	99%	99%	56%	79%

...

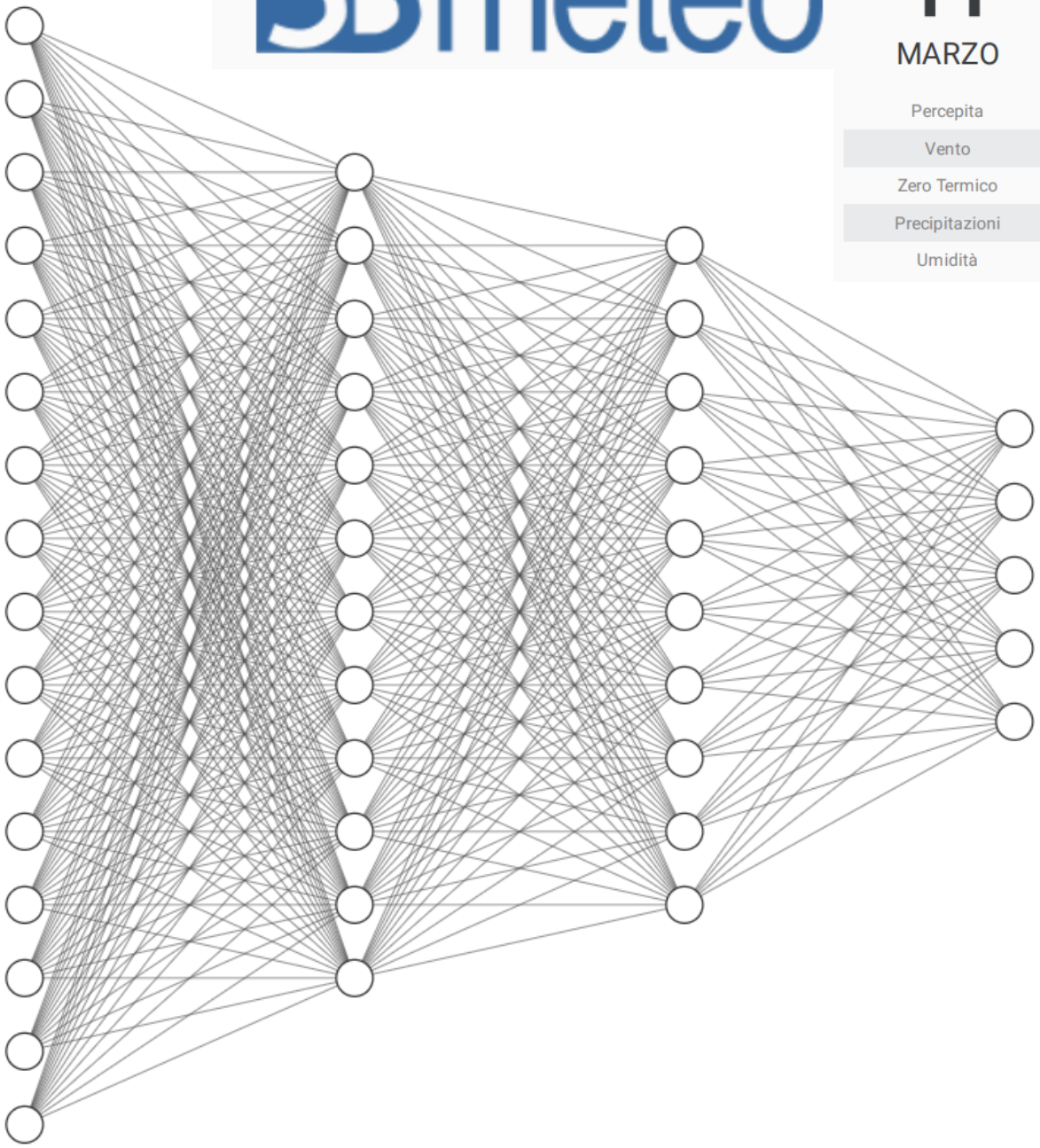
Temperatura

Umidità

Vento

Tipologia Nuvole

...



Precipitazione %

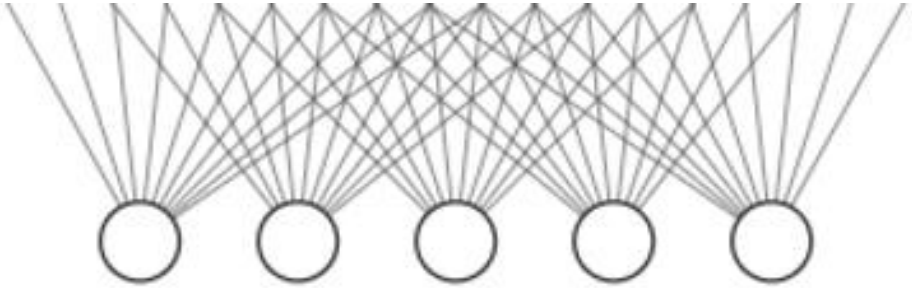
T massima

T minima

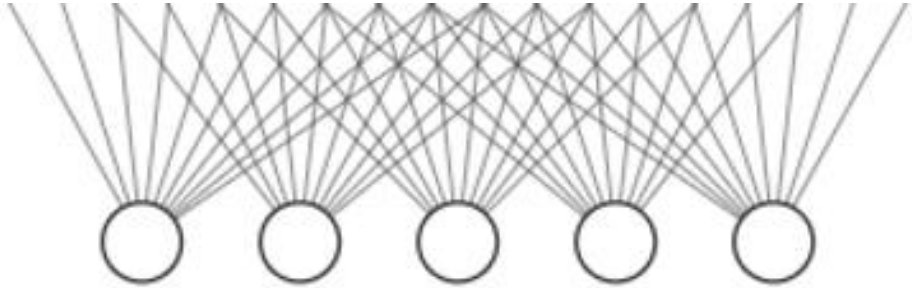
UV

...

Errore / Loss



0.8 15 12 4 5



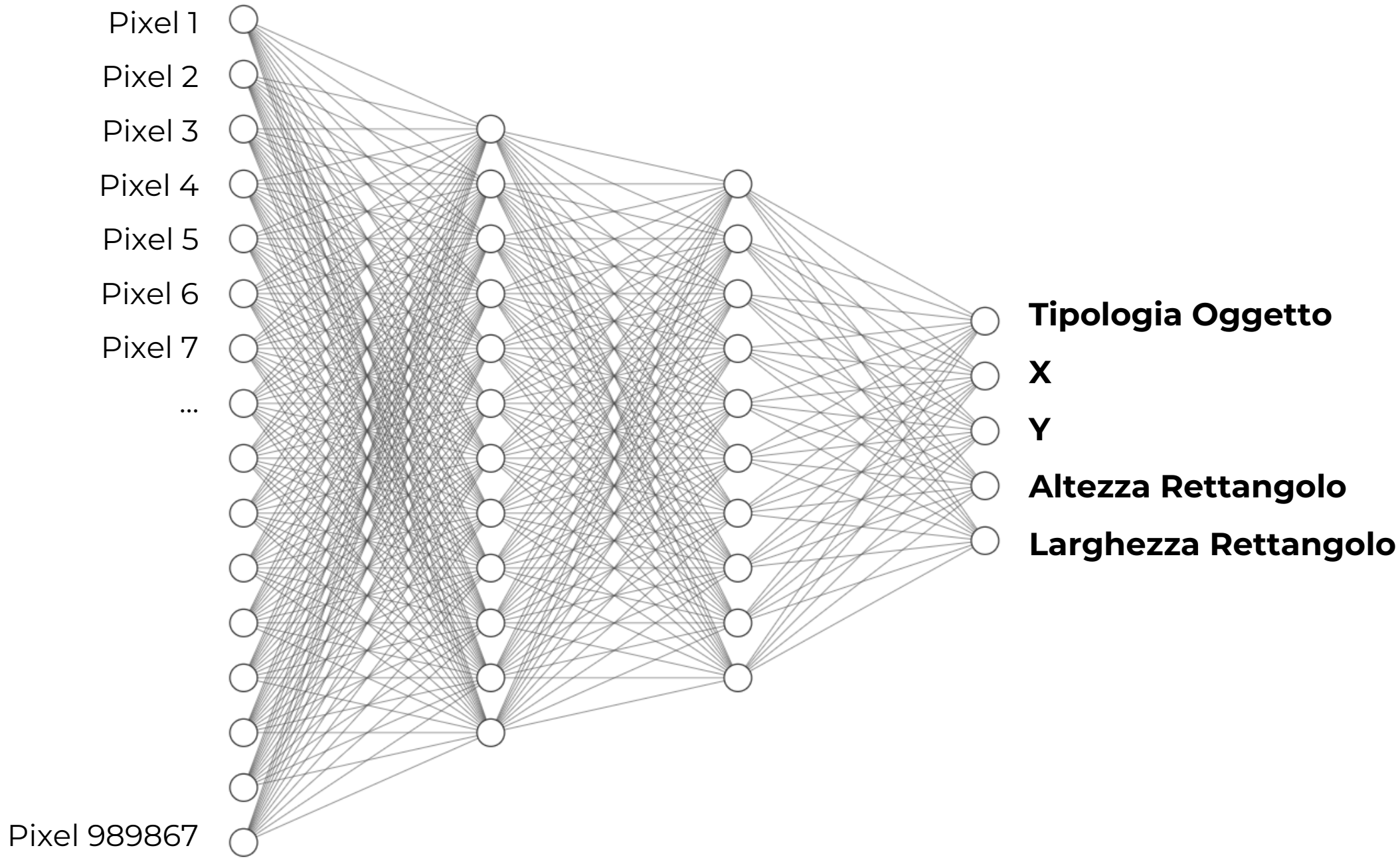
0.7 13 6 4 0

$$MSE = \frac{1}{N}\sum_{i-1}^{i=N}(y_i - \widehat{y_i})^2$$

$$MSE = \frac{1}{3}[(y_1 - \widehat{y_1})^2 + (y_2 - \widehat{y_2})^2 + (y_3 - \widehat{y_3})^2]$$

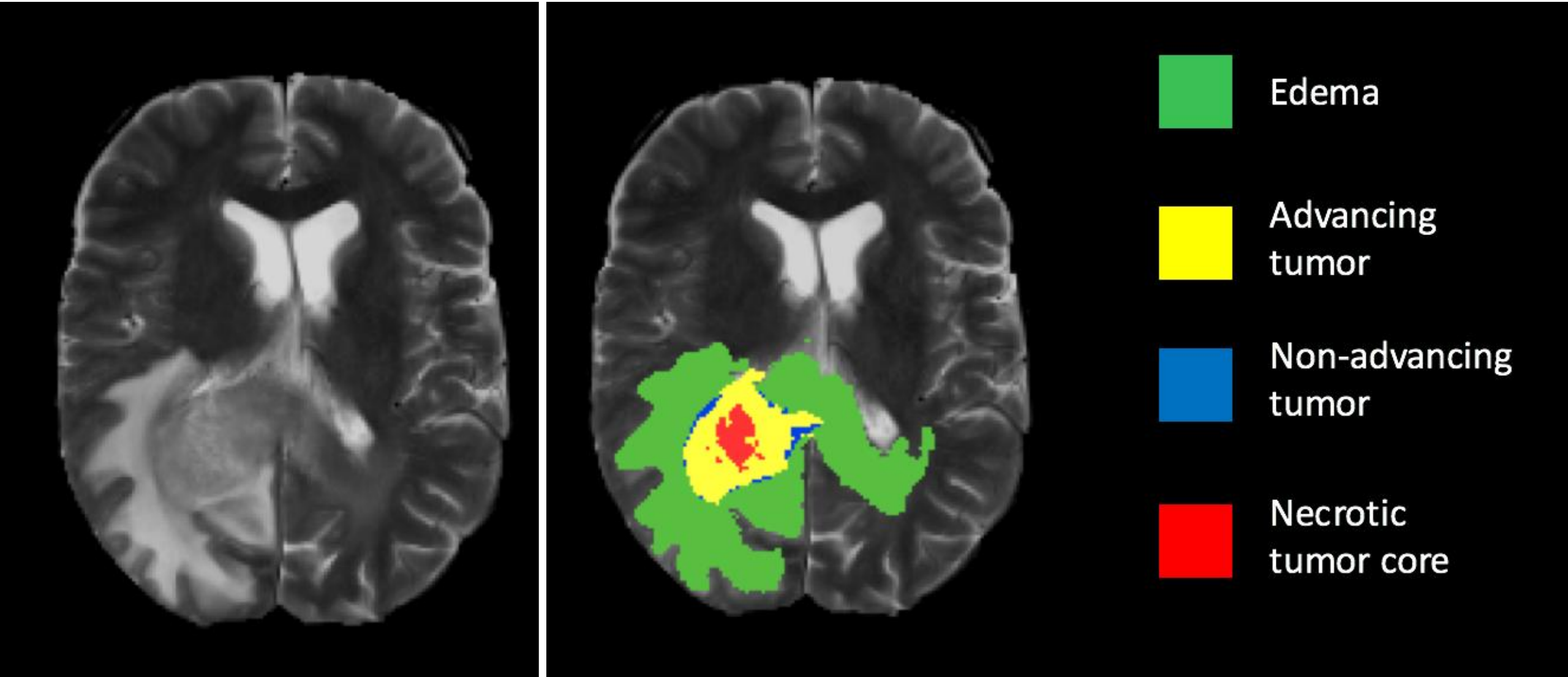
Detection

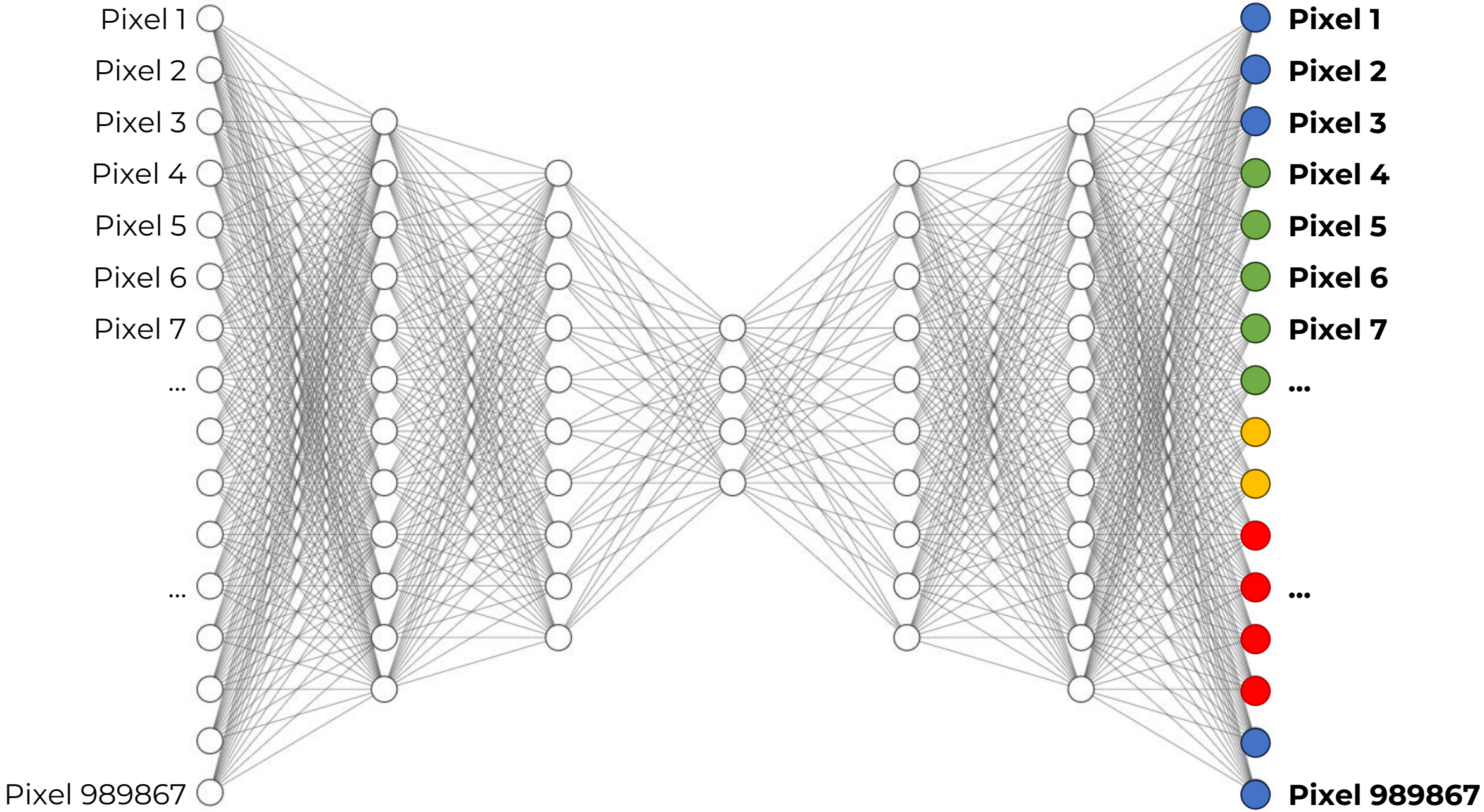
Dove è quello che cerco? (approssimativamente)


Pixel 1
Pixel 2
Pixel 3
Pixel 4
Pixel 5
Pixel 6
Pixel 7
...
Pixel 989867
Tipologia Oggetto
X
Y
Altezza Rettangolo
Larghezza Rettangolo

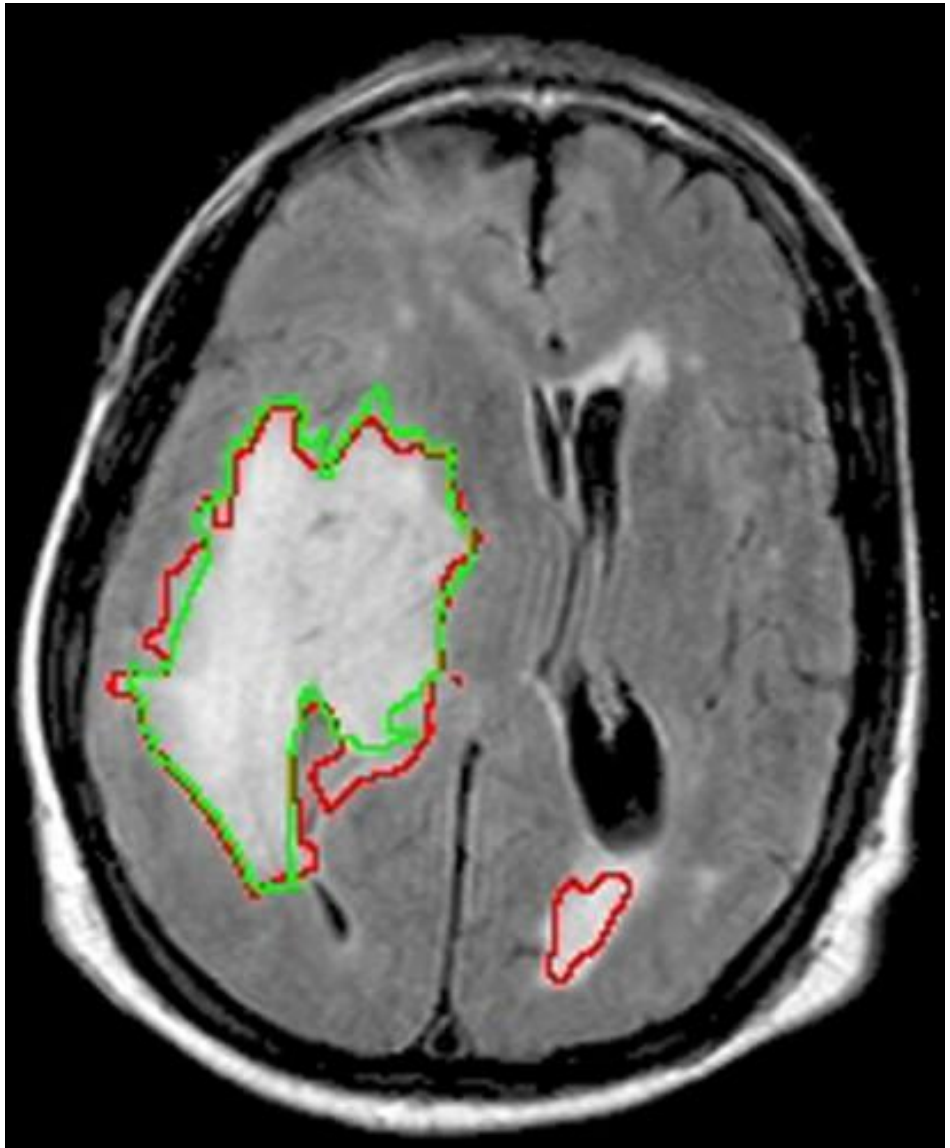
Segmentation

Dove è quello che cerco? (precisamente)

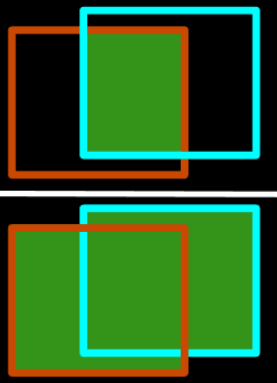



Pixel 1
Pixel 2
Pixel 3
Pixel 4
Pixel 5
Pixel 6
Pixel 7
...
...
Pixel 989867
Pixel 1
Pixel 2
Pixel 3
Pixel 4
Pixel 5
Pixel 6
Pixel 7
...
...
Pixel 989867

Errore / Loss



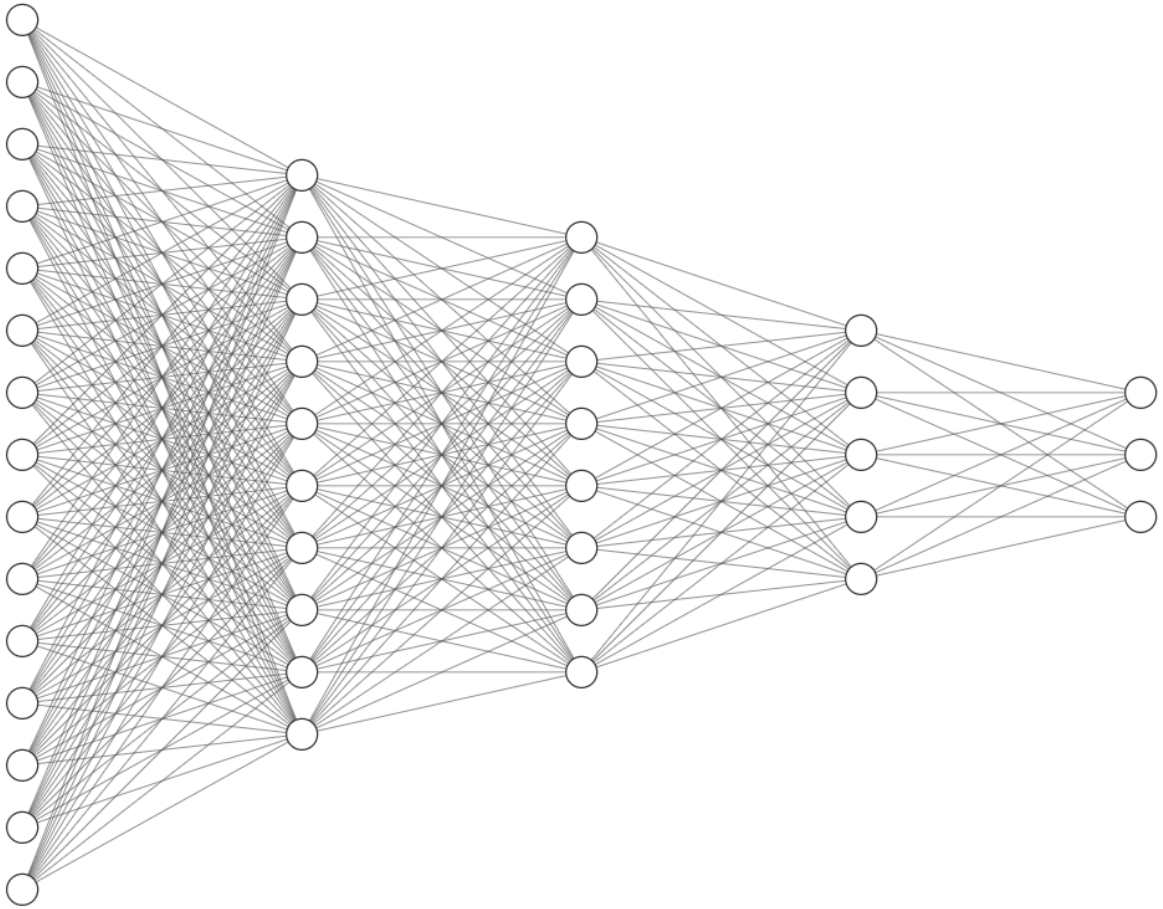
$$IOU = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$



Generation

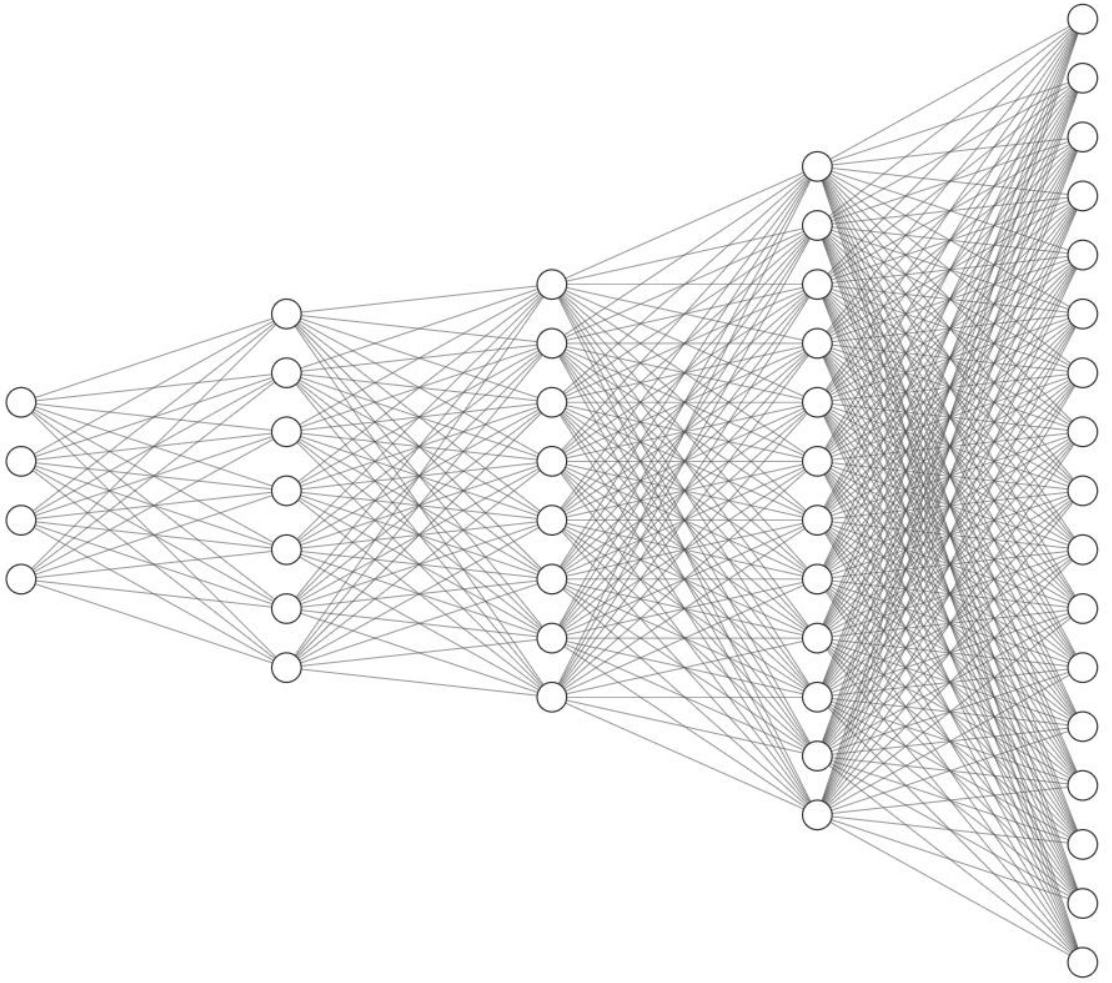
“Pope Francis swagging in a white puffer jacket“

Discriminative AI

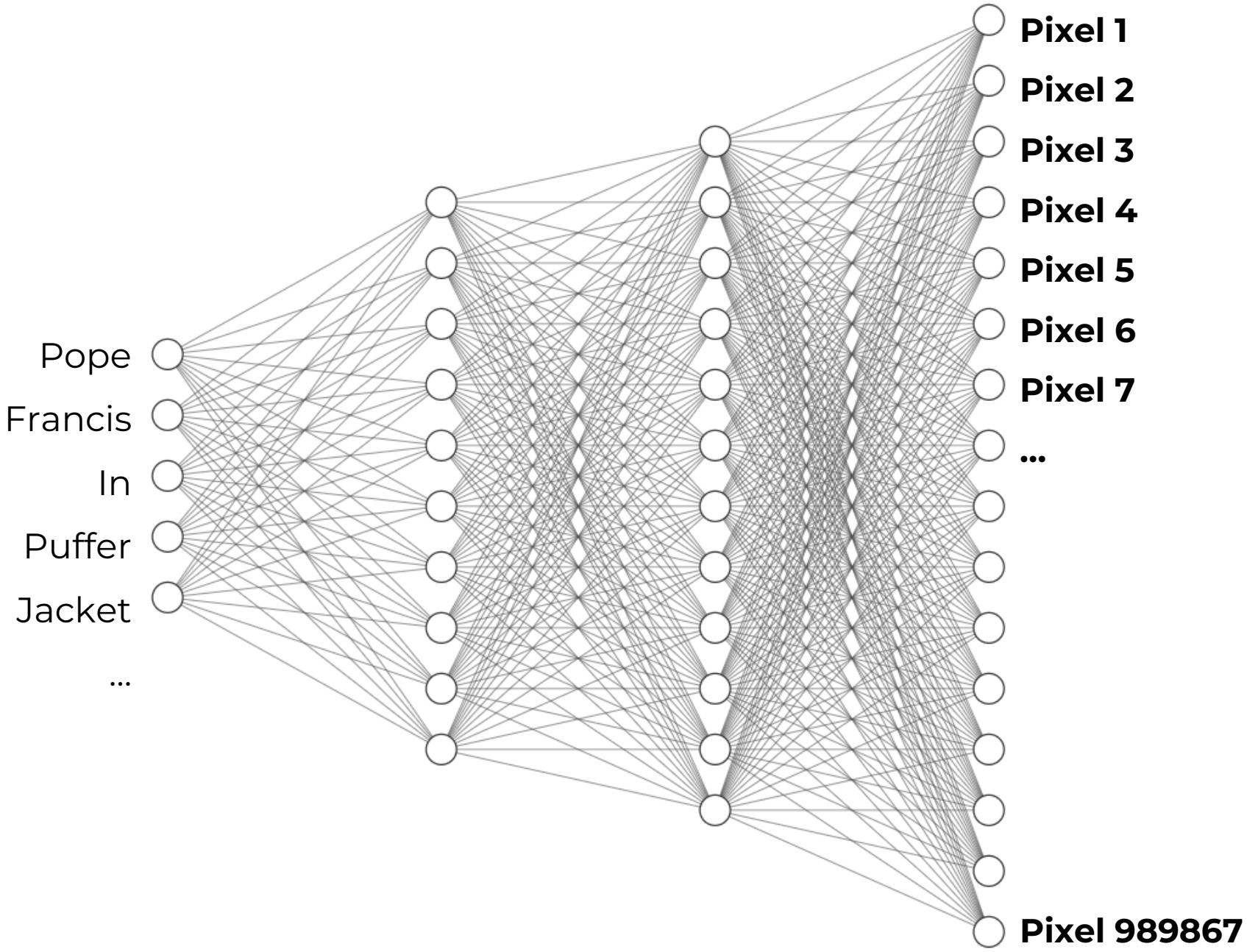


Big Input→ Small Output

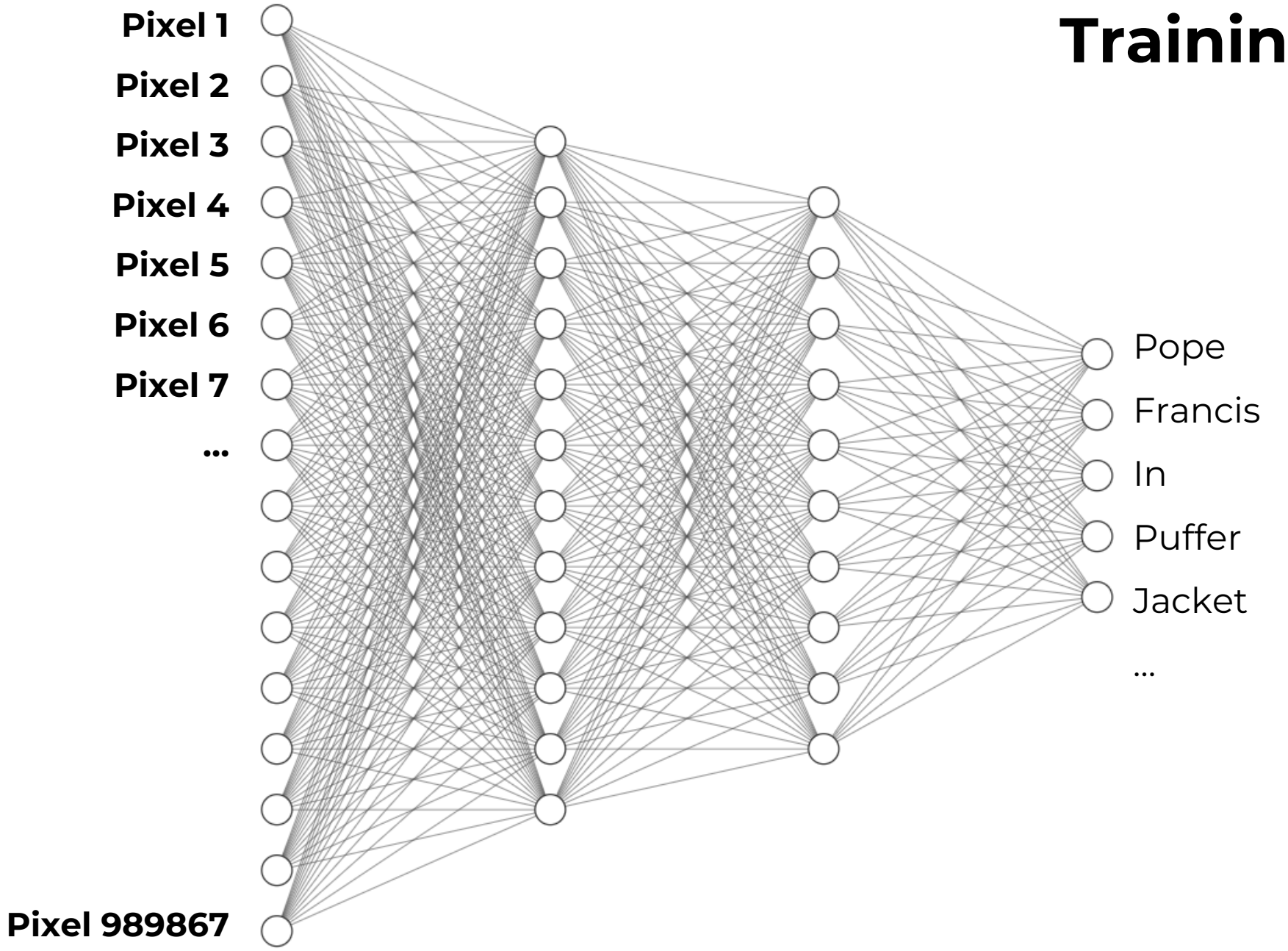
Generative AI



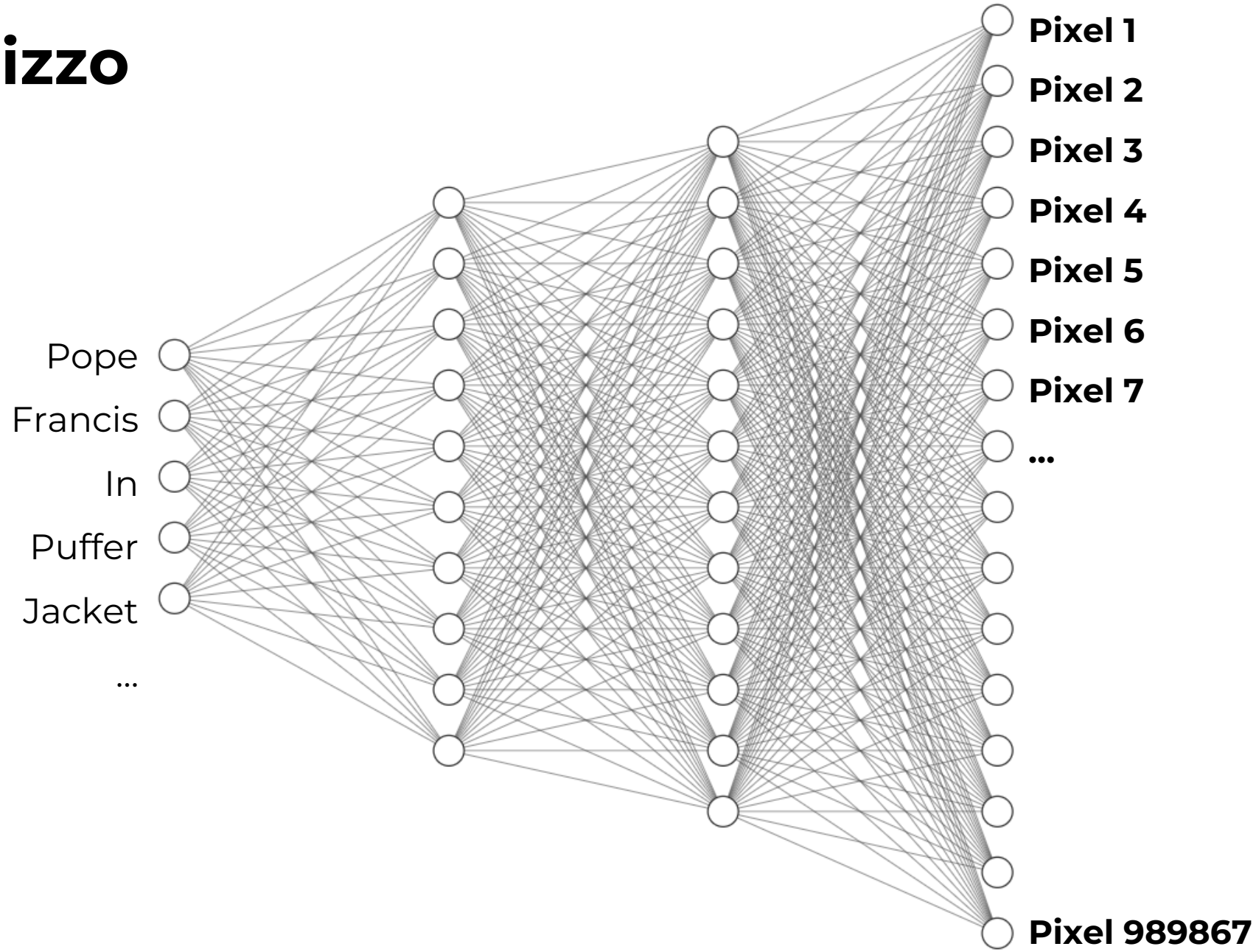
Small Input → Big Output


Pope
Francis
In
Puffer
Jacket
...
Pixel 1
Pixel 2
Pixel 3
Pixel 4
Pixel 5
Pixel 6
Pixel 7
...
Pixel 989867

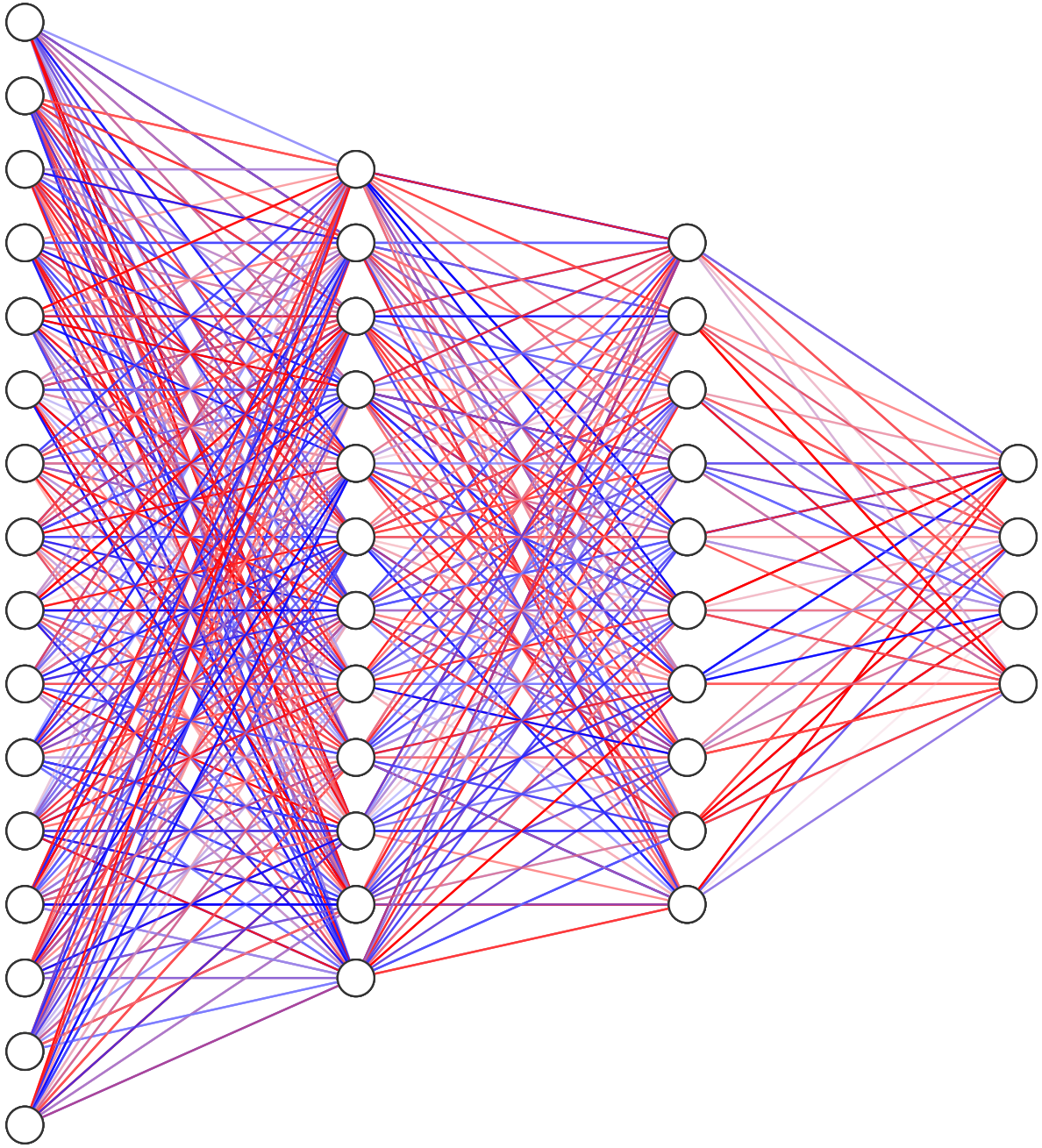
Training

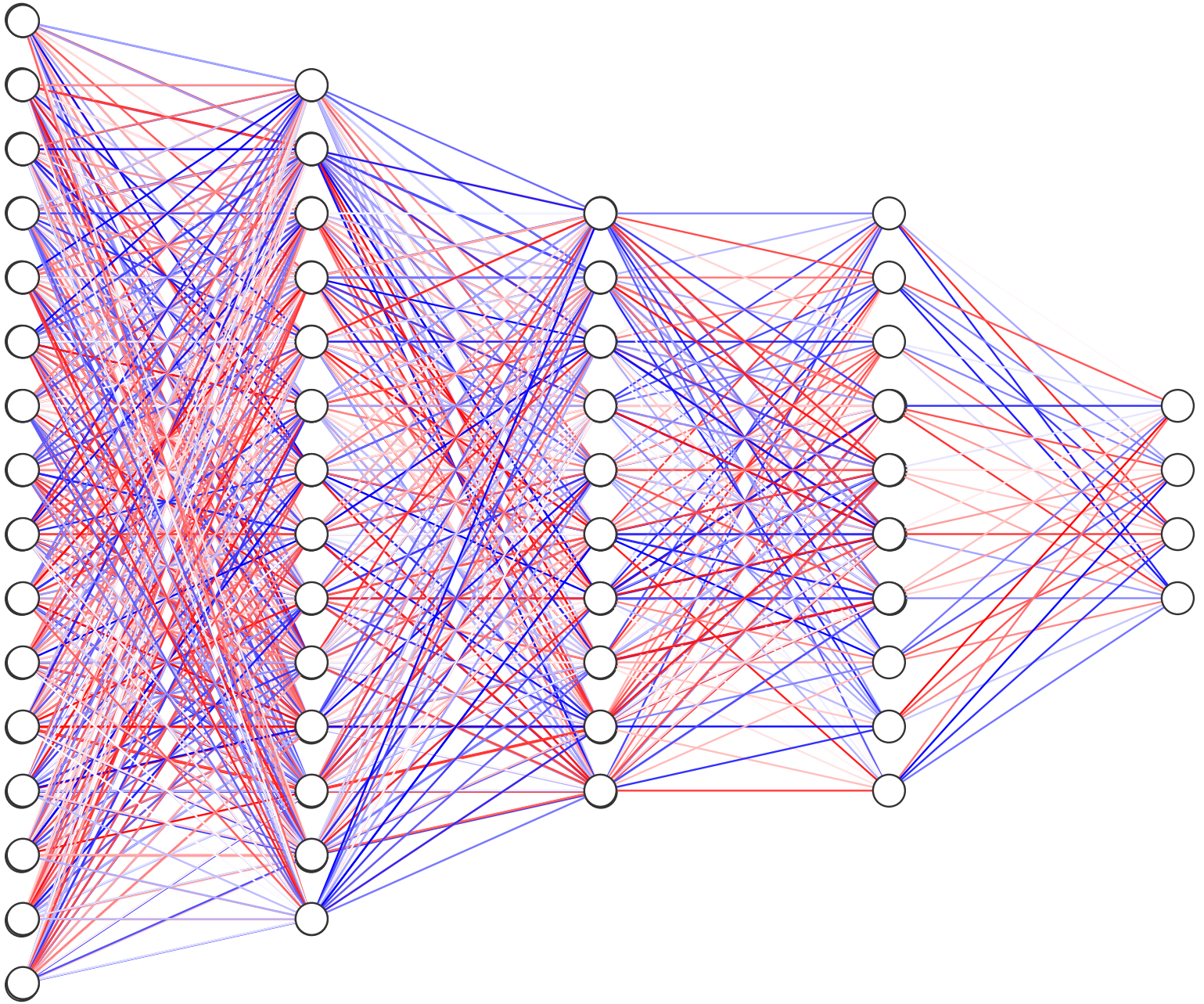

Pixel 1
Pixel 2
Pixel 3
Pixel 4
Pixel 5
Pixel 6
Pixel 7
...
Pixel 989867
Pope
Francis
In
Puffer
Jacket
...

Utilizzo


Pope
Francis
In
Puffer
Jacket
...
Pixel 1
Pixel 2
Pixel 3
Pixel 4
Pixel 5
Pixel 6
Pixel 7
...
Pixel 989867

HyperParameters





Parameters

- **Internal** to the model
- **Estimated** from the data
- **Required** by the model
- **Not set manually** by the practitioner

Example of model parameters:

- The **weights**

HyperParameters

- **External** to the model
- **Cannot be estimated** from the data
- **Tuned** for model predictions
- **Set manually** by the practitioner.

Examples of model parameters:

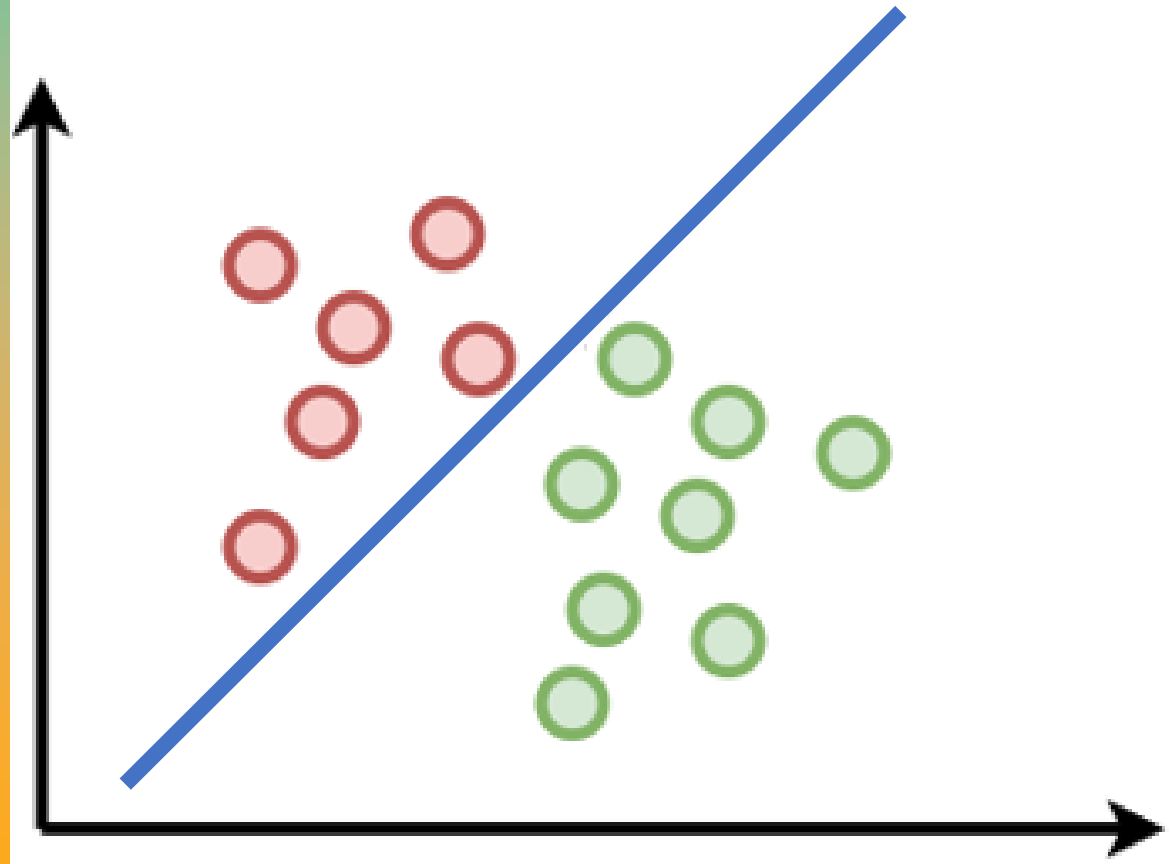
- The **learning rate**
- The **number of hidden layers**
- The **batch size**

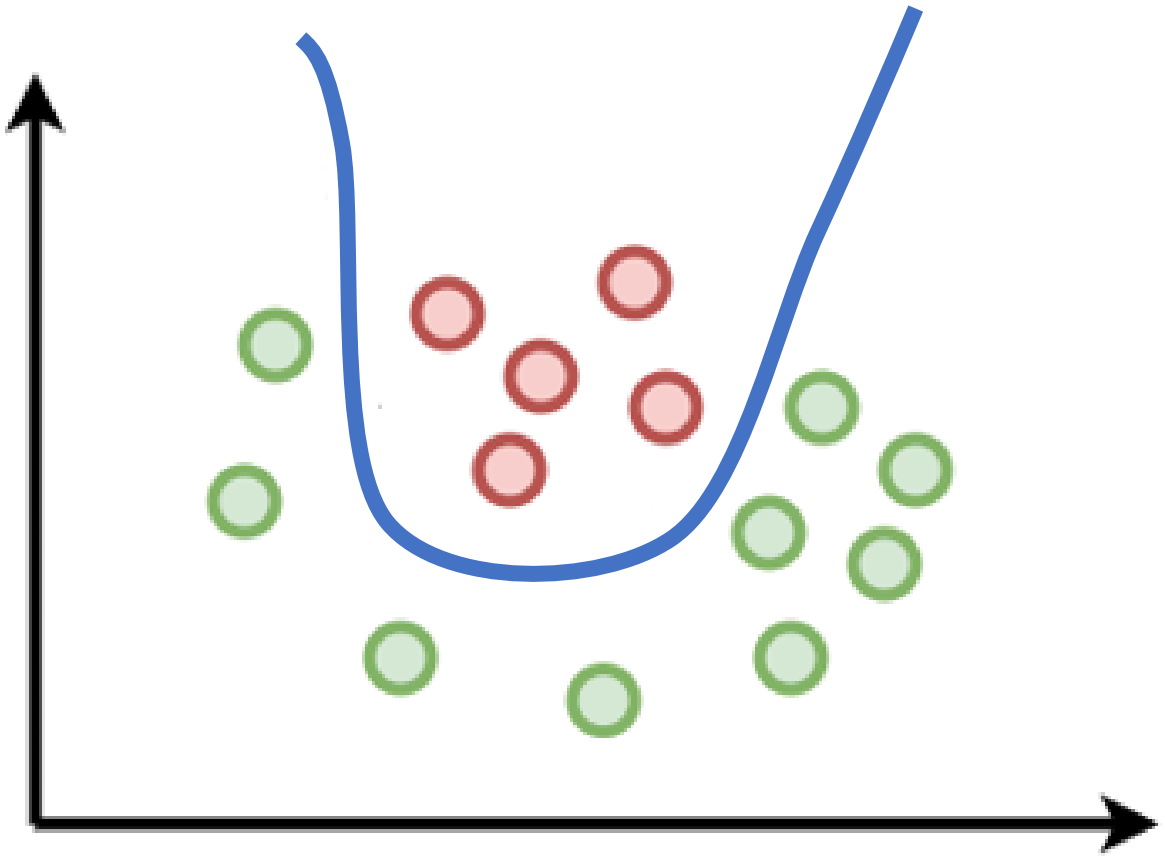
Activation Function

$$y = x_1 \cdot w_1 + x_2 \cdot w_2 + x_3 \cdot w_3 + x_4 \cdot w_4 + x_5 \cdot w_5$$

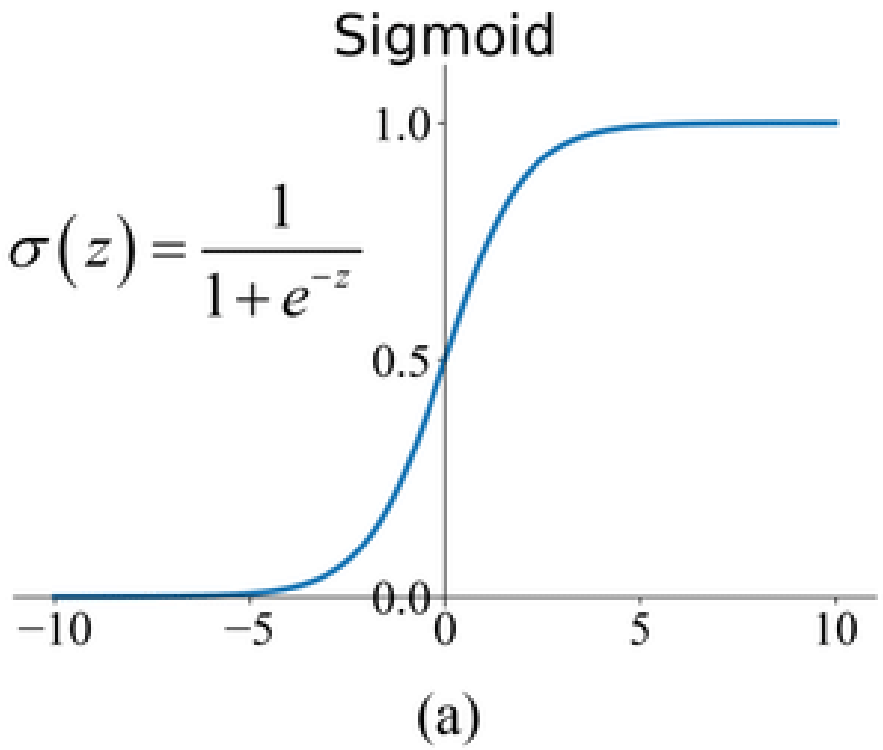
$$z = x_1 \cdot w_1 + x_2 \cdot w_2 + x_3 \cdot w_3 + x_4 \cdot w_4 + x_5 \cdot w_5$$

$$y = \frac{1}{1 + e^{-z}} = \frac{1}{1 + e^{-(x_1 \cdot w_1 + x_2 \cdot w_2 + x_3 \cdot w_3 + x_4 \cdot w_4 + x_5 \cdot w_5)}}$$

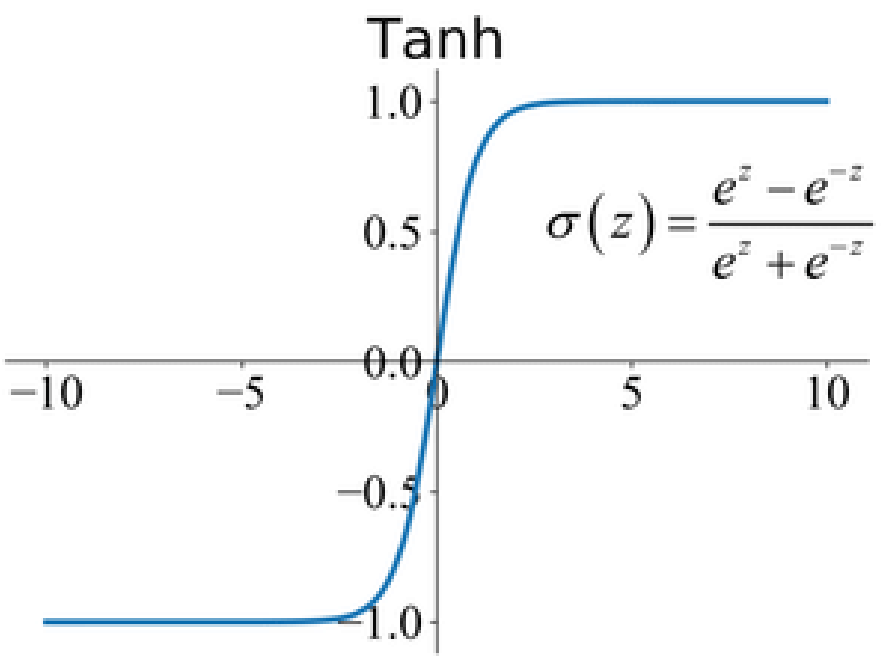




Activation Functions



(a)

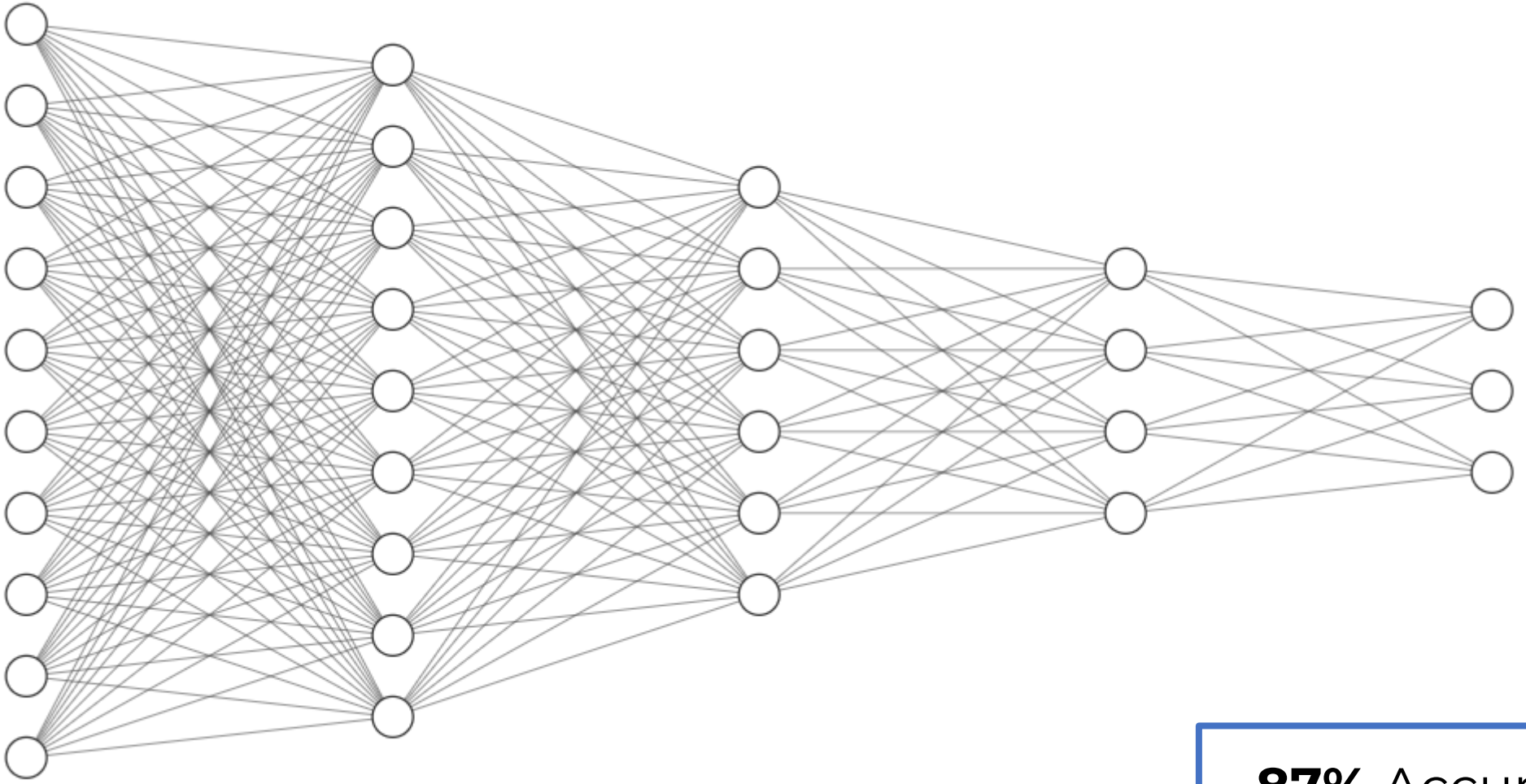


(b)

Overfitting
e
Generalizzazione

10

INPUT



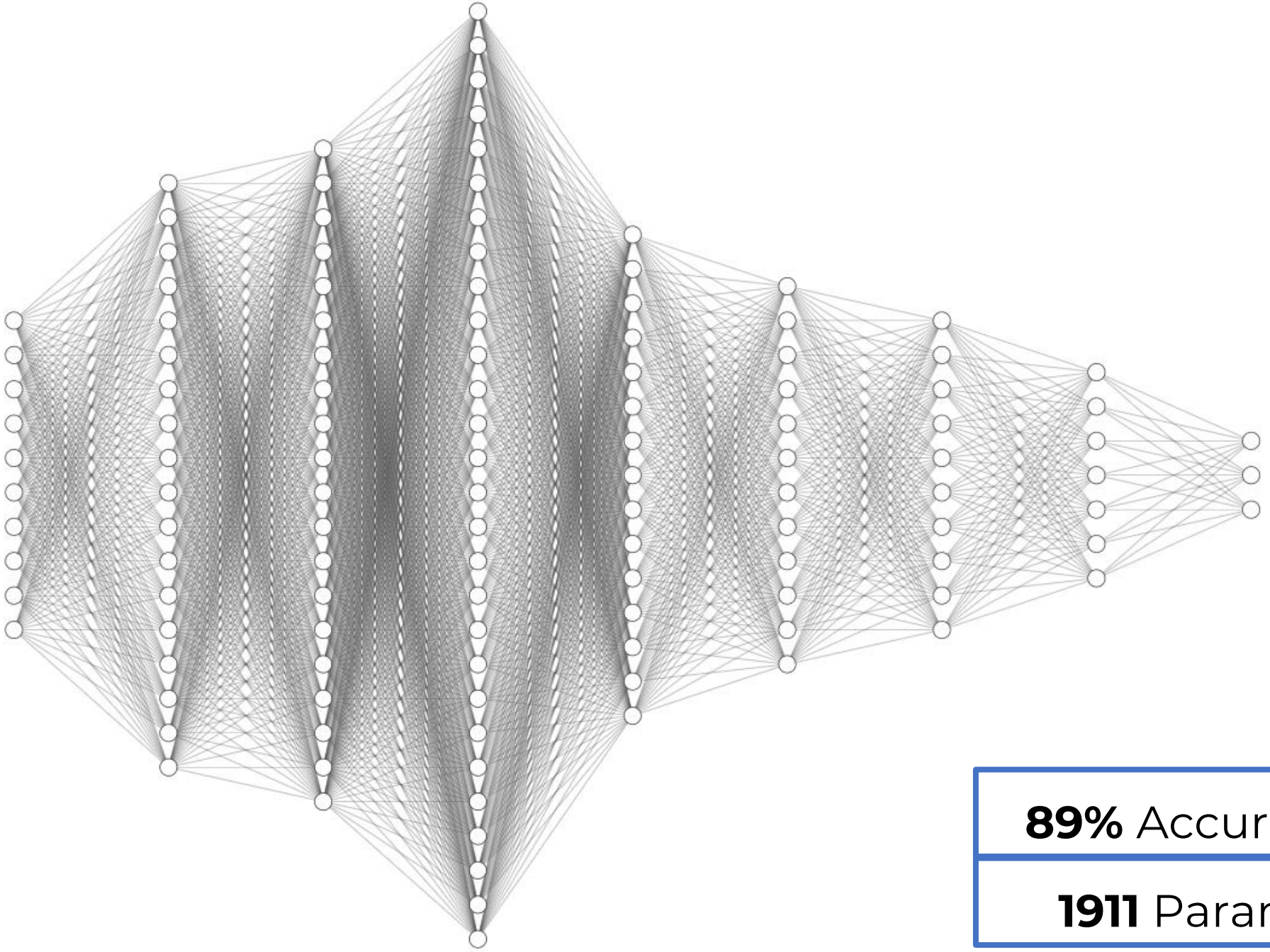
3

OUTPUT

87% Accuratezza
180 Parametri

10

INPUT



3

OUTPUT

89% Accuratezza

1911 Parametri

$(1+3)^3 = ?$

$$(1+3)^3 =$$
$$= 4^3$$
$$= \mathbf{64}$$

$$(A+B)^3 =$$

$$= (A+B)(A+B)(A+B) =$$

$$= (A^2 + AB + AB + B^2)(A+B) =$$

$$= (A^2 + 2AB + B^2)(A+B) =$$

$$= A^3 + A^2B + 2A^2B + 2AB^2 + AB^2 + B^3 =$$

$$= A^3 + 3A^2B + 3AB^2 + B^3$$

$$= \mathbf{64}$$

Dataset Splitting

Imparo con: 70%

Controllo con: 30%

abc Cust City	abc Cust State	abc Cust Zip Code	abc Cust County	abc Cust Region
Costa Mesa	California	92628	Orange County, California	West
Edinburg	Texas	78539	Hidalgo County, Texas	Southwest
Vancouver	Washington	98668	Clark County, Washington	West
El Paso	Texas	79910	El Paso County, Texas	Southwest
Columbus	Georgia	31908	Muscogee County, Georgia	South
Omaha	Nebraska	68108	Douglas County, Nebraska	Midwest
Mountain View	California	94035	Contra Costa County, Californ	West
Lincoln	Nebraska	68501	Lancaster County, Nebraska	Midwest
Tulsa	Oklahoma	74107	Osage County, Oklahoma	Southwest
Gastonia	North Carolina	28052	Gaston County, North Carolin	South
Newport News	Virginia	23501	Newport News City, Virginia	East
Dayton	Ohio	45401	Montgomery County, Ohio	Midwest
South Gate	California	90280	Los Angeles County, Californi	West
Birmingham	Alabama	35203	Jefferson County, Alabama	South
Sacramento	California	95813	Sacramento County, Californi	West
Dallas	Texas	75260	Dallas County, Texas	Southwest
Antioch	California	94509	Contra Costa County, Californ	West
Lubbock	Texas	79402	Lubbock County, Texas	Southwest
Bellflower	California	90706	Los Angeles County, Californi	West
St. Paul	Minnesota	55109	Ramsey County, Minnesota	Midwest
Mesquite	Texas	75149	Dallas County, Texas	Southwest
Rio Rancho	New Mexico	87124	Bernalillo County, New Mexic	Southwest
Edinburg	Texas	78539	Hidalgo County, Texas	Southwest
Camden	New Jersey	08101	Camden County, New Jersey	East
Sioux Falls	South Dakota	57104	Lincoln County, South Dakota	Midwest
Lexington-Fayette	Kentucky	40511	Fayette County, Kentucky	South
Louisville	Kentucky	40231	Jefferson County, Kentucky	South
Fall River	Massachusetts	02720	Bristol County, Massachusett	East
Palmdale	California	93550	Los Angeles County, Californi	West
Sacramento	California	95813	Sacramento County, Californi	West
Sioux Falls	South Dakota	57104	Lincoln County, South Dakota	Midwest
Austin	Texas	78710	Travis County, Texas	Southwest
Quincy	Massachusetts	02169	Norfolk County, Massachuset	East
Roswell	Georgia	30075	Fulton County, Georgia	South
Oakland	California	94612	Alameda County, California	West
Bellflower	California	90706	Los Angeles County, Californi	West
Phoenix	Arizona	85026	Maricopa County, Arizona	Southwest

Quanti neuroni?
Quanti layers?
Quale Loss?
Quale LR?
Quale Batch Size?
Quale attivazione?