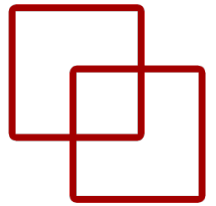
Improving Deep Learning by Exploiting Synthetic Images

UNICTEC

Dr. Manuel Castillo-Cara
www.manuelcastillo.eu
Departamento de Inteligencia Artificial
Escuela Técnica Superior de Ingeniería Informática
Universidad Nacional de Educación a Distancia (UNED)

Preliminar

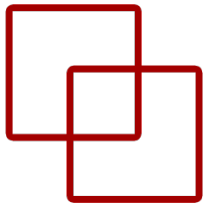


-

UNED

ETS de Ingeniería Informática

Índice



- Estudiar en la UNED
- Ciencia de datos
- Datos… ¿Por qué ahora?
- Avances del Deep Learning
- Arquitecturas neuronales
- Fundamentos de imágenes sintéticas
- Métodos de transformación
- Arquitecturas con imágenes sintéticas
- Casos de uso

Curriculum Vitae

ETS de Ingeniería Informática

UNED

Almería

Imágenes tomadas de diferentes sitios de internet para una finalidad interna y no pública

Almería

Nicolás Salmerón fotografiado por Compañy, publicado en la revista *Nuevo Mundo* el 24 de septiembre de 1908.

Presidente del Poder Ejecutivo de la República Española

18 de julio de 1873-7 de septiembre de 1873

Predecesor	Francisco Pi y Margall
Sucesor	Emilio Castelar

Imágenes tomadas de diferentes sitios de internet para una finalidad interna y no pública

Almería

ás Salmerón fotografiado por Compañy,
ublicado en la revista *Nuevo Mundo*
el 24 de septiembre de 1908.

dente del Poder Ejecutivo de la
República Española
io de 1873-7 de septiembre de 1873
or Francisco Pi y Margall
Sucesor Emilio Castelar

Imágenes tomadas de diferentes sitios de internet para una finalidad interna y no pública

Doctorado – UCLM Albacete

Imágenes tomadas de diferentes sitios de internet para una finalidad interna y no pública

UNI

Imágenes tomadas de diferentes sitios de internet para una finalidad interna y no pública

IUT–SCi

Imágenes tomadas de diferentes sitios de internet para una finalidad interna y no pública

ULima

Imágenes tomadas de diferentes sitios de internet para una finalidad interna y no pública

OEG – UPM


jadahonda
EL PLANTÍO
Hospital Universitario Puerta De Hierro
Pozuelo de Alarcón
ARAVACA
Escuela Técnica Superior de Ingenieros...
EL PILAR
CHAMARTÍN
Escuela Técnica Superior de Ingeniería...
HORTALEZA
Terminal T2 Salidas
BARAJAS
Centro comercial Arturo Soria Plaza
CDAD. LINEAL
Parroquia Santa María de Caná
Escuela Técnica Superior de Edificación UPM
LA PAZ
Húmera
Casa de Campo
SALAMANCA
Parque Quinta de la Fuente del Berro
CLUB LAS ENCINAS DE BOADILLA
Somosaguas
Madrid
RETIRO
Universidad Rey Juan Carlos (URJC) Campus de...
Escuela Técnica Superior de Ingenieros...
LATINA
ATOCHA
Cerro del Tío Pío
Monteprincipe
Matadero Madrid
Universidad Politécnica de Madrid: Escuela...
PUENTE DE VALLECAS
VISTA ALEGRE
USERA
SANTA EUGENIA
ENTREVÍAS
CARABANCHEL
Venta la Rubia
Parque de La Gavia
Sambil Outlet Madrid
BUENAVISTA
ORCASITAS
VALLECAS


UPM

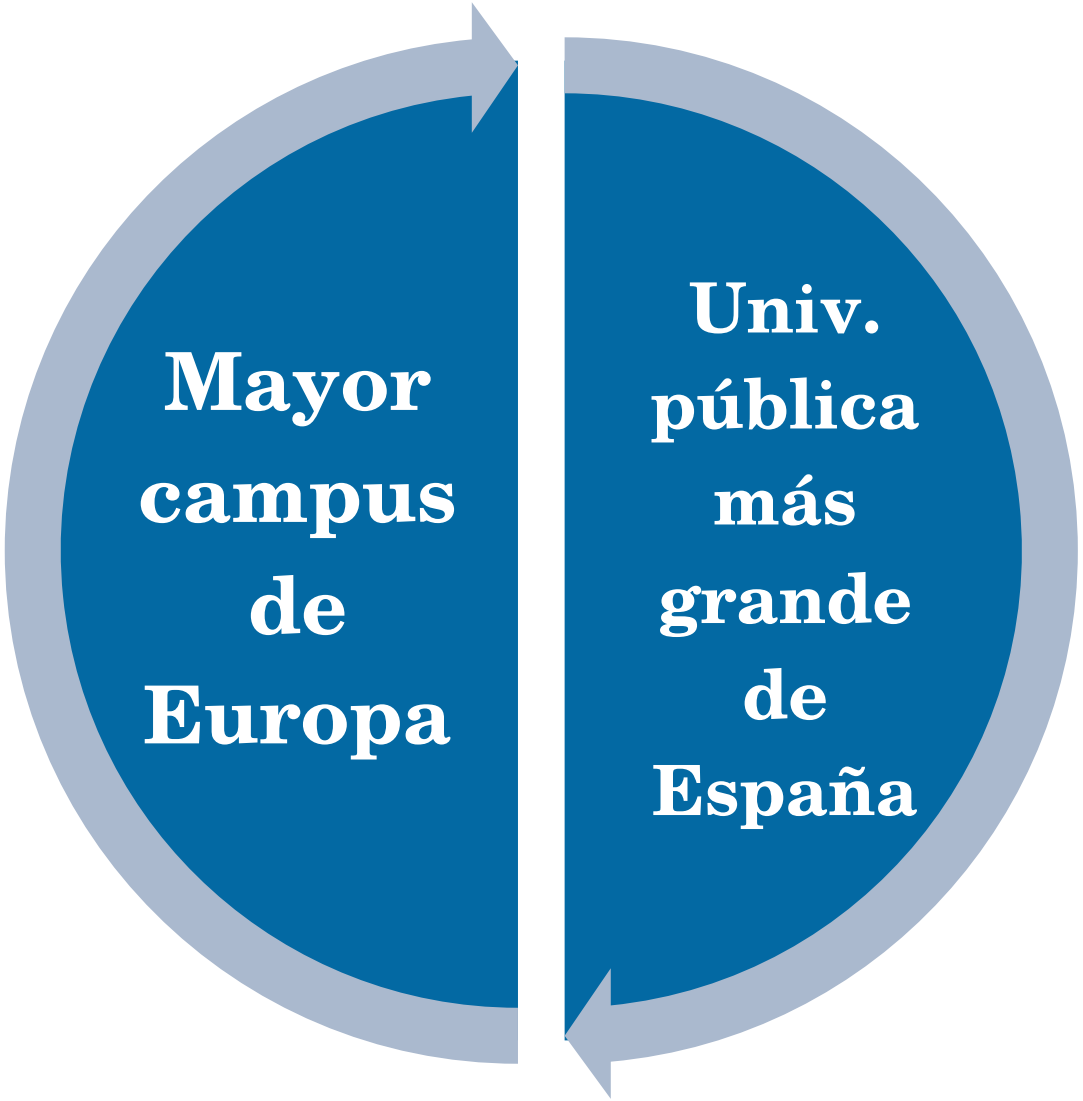
UNED

Estudiar en la UNED

ETS de Ingeniería Informática

UNED

La UNED



UNED en números

+ 30 grados

+20 grados combinados

+80 Másteres EEES

20 Programas Doctorado Internacional

+16 microtítulos

+800 títulos propios

+500 cursos online gratuitos

+3’5 millones de € en contratos de Transferencia

+150.000 estudiantes en enseñanzas regladas

+3.000.000 alumni

+100 sedes UNED

17 países

Dimensión Geográfica Internacional

América

Centros de la UNED

- Centro en Buenos Aires
- Centro en Caracas
- Centro en Lima
- Centro en México
- Centro en Sao Paulo

Aula de examen de la UNED

- Aula en Bogotá
- Aula en Nueva York
- Aula en Quito
- Aula en Santiago de Chile
- Aula en Seattle

Europa

Centros de la UNED

- Centro en Berlín
- Centro en Berna
- Centro en Bruselas
- Centro en Gjirokaster (Albania)
- Centro en Lisboa
- Centro en Londres
- Centro en París

Aula de examen de la UNED

- Aula en Frankfurt
- Aula en Munich
- Aula en Oporto
- Aula en Roma
- Aula en Varsovia y Europa del este

África

Centros de la UNED

- Centro en Bata
- Centro en Malabo

UNED – Perú

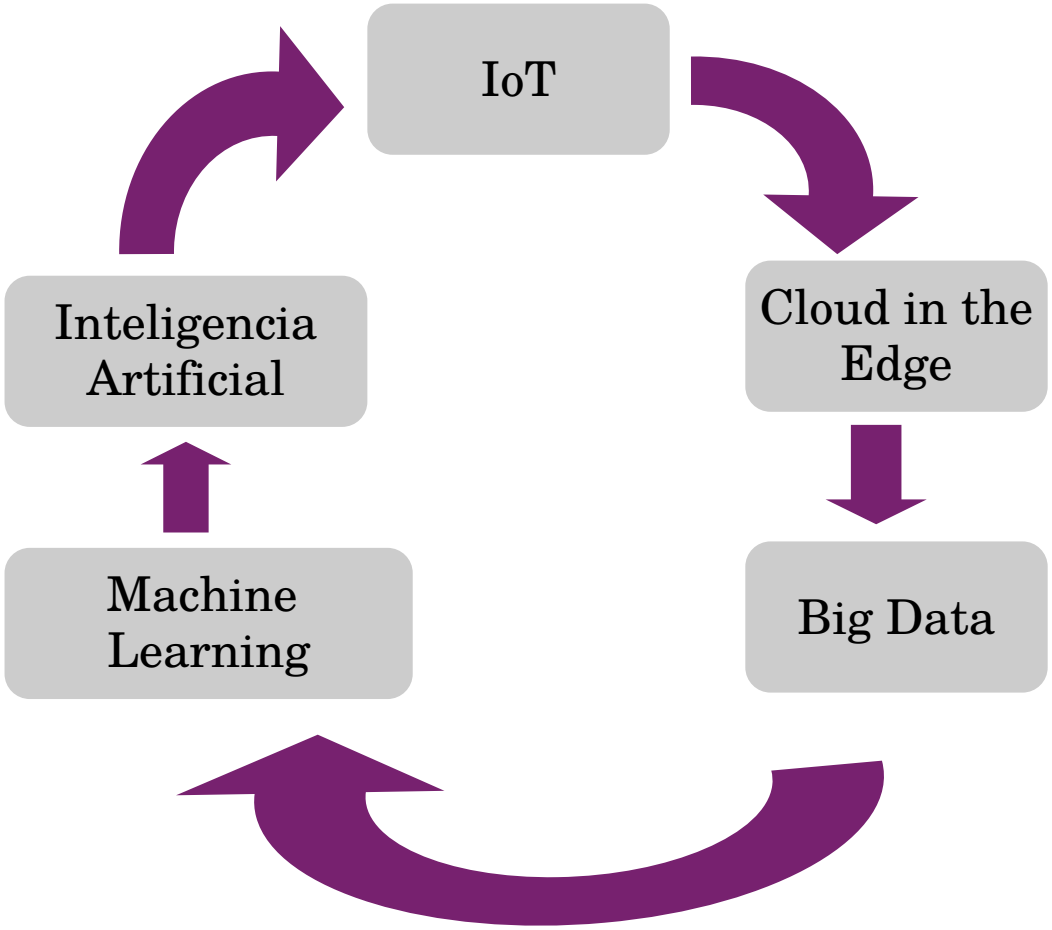
- Universidad pública con representación internacional.
  - Tasas muy bajas con una exigencia alta.
  - Títulos europeos oficiales
- En Lima, hay centro UNED.
- Centro Español en Salaverry. Ubicación.
- Posibilidad de estudiar pregrado y posgrado:
  - Máster de Inteligencia Artificial
  - Máster en Ciencia e Ingeniería de Datos
  - Más masters en diferentes áreas.

Líneas de investigación

ETS de Ingeniería Informática

UNED

ICBM-AI



Investigación, desarrollo e innovación (I+D+i)

Ciencia de datos

ETS de Ingeniería Informática

UNED

Noticias

Armas digitales contra el fraude económico

Big Data contra la pesca ilegal

Grandes oportunidades para avanzar en la investigación sanitaria gracias al 'big data'

Un futuro alimentario en manos de la tecnología (pero no solo)

Facebook desarrolla un algoritmo para detectar conductas suicidas antes que los humanos

Noticias

Armas digitales contra el fraude económico

Big Data contra la pesca ilegal

Grandes oportunidades para avanzar en la investigación sanitaria gracias al 'big data'

Un futuro alimentario en manos de la tecnología (pero no solo)

Facebook desarrolla un algoritmo para detectar conductas suicidas antes que los humanos

El mundo de los datos necesita ingenieros

La fiebre del 'Big Data' llega al empleo

Noticias

Armas digitales contra el fraude económico

Big Data contra la pesca ilegal

Grandes oportunidades para avanzar en la investigación sanitaria gracias al 'big data'

Un futuro alimentario en manos de la tecnología (pero no solo)

Facebook desarrolla un algoritmo para detectar conductas suicidas antes que los humanos

El mundo de los datos necesita ingenieros

La fiebre del 'Big Data' llega al empleo

La guerra por nuestros datos

La era del algoritmo ha llegado y tus datos son un tesoro

"La mayor parte de las veces que nos analizan en Internet ni nos enteramos"

Noticias

Armas digitales contra el fraude económico

Big Data contra la pesca ilegal

Grandes oportunidades para avanzar en la investigación sanitaria gracias al 'big data'

Un futuro alimentario en manos de la tecnología (pero no solo)

Facebook desarrolla un algoritmo para detectar conductas suicidas antes que los humanos

El mundo de los datos necesita ingenieros

La fiebre del 'Big Data' llega al empleo

Algoritmo: ángel o demonio

Los salarios en Google los programa un algoritmo

"Los algoritmos reproducen las desigualdades del mundo real"

Google arregla su algoritmo 'racista' borrando a los gorilas

La guerra por nuestros datos

La era del algoritmo ha llegado y tus datos son un tesoro

"La mayor parte de las veces que nos analizan en Internet ni nos enteramos"

Noticias

Armas digitales contra el fraude económico

Big Data contra la pesca ilegal

Grandes oportunidades para avanzar en la investigación sanitaria gracias al 'big data'

Se rompe la fe ciega en el Big Data

El 'big data' es el pulpo Paul del Mundial de Rusia 2018 y dice que España será subcampeona

Un futuro alimentario en manos de la tecnología (pero no solo)

Facebook desarrolla un algoritmo para detectar conductas suicidas antes que los humanos

Cambridge Analytica ofreció sus servicios en España

El algoritmo que 'adivina' los delitos futuros falla tanto como un humano

El mundo de los datos necesita ingenieros

La fiebre del 'Big Data' llega al empleo

Algoritmo: ángel o demonio

Los salarios en Google los programa un algoritmo

"Los algoritmos reproducen las desigualdades del mundo real"

Google arregla su algoritmo 'racista' borrando a los gorilas

La guerra por nuestros datos

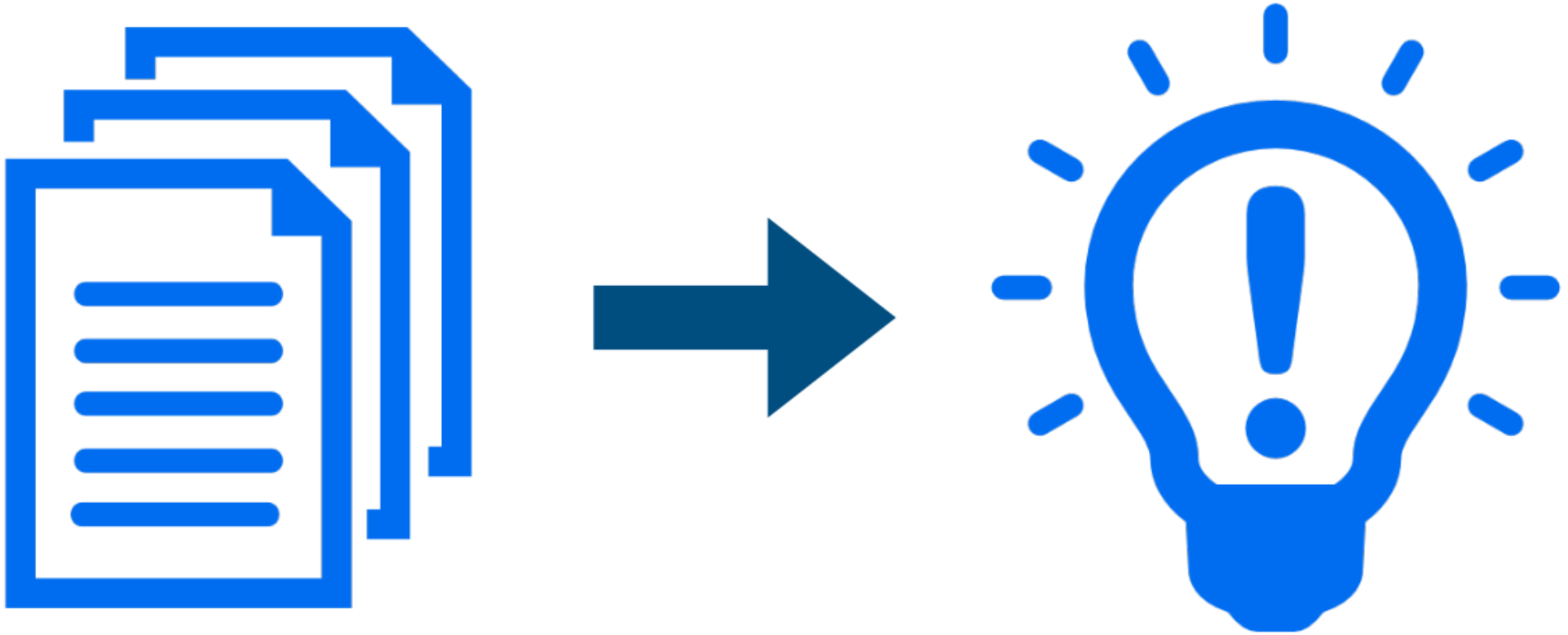
La era del algoritmo ha llegado y tus datos son un tesoro

"La mayor parte de las veces que nos analizan en Internet ni nos enteramos"

Objetivo común

Datos

Objetivo común



Datos

Conocimiento

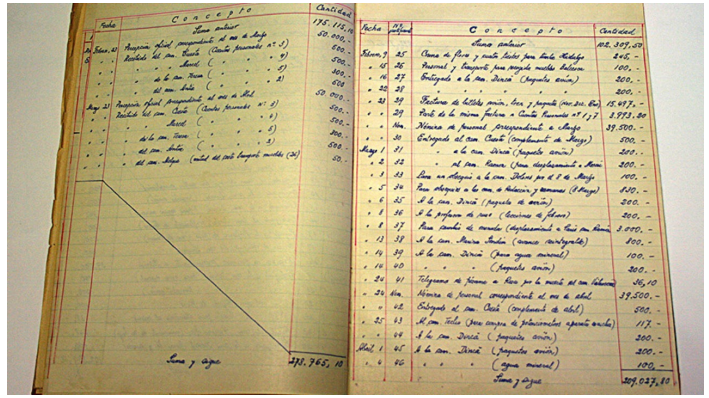
Datos... ¿Por qué ahora?

ETS de Ingeniería Informática

UNED

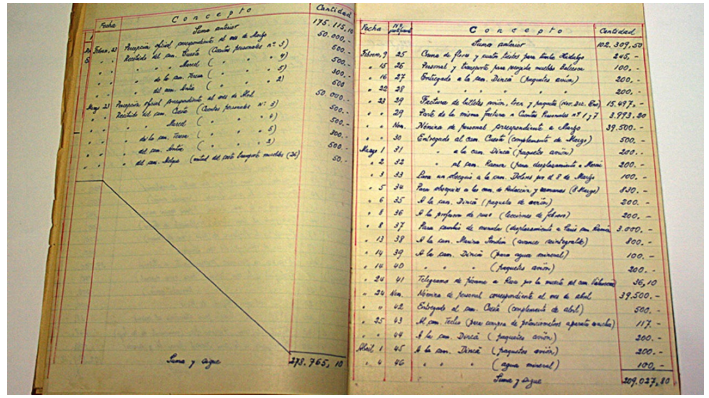
Avances

- Producción de datos.



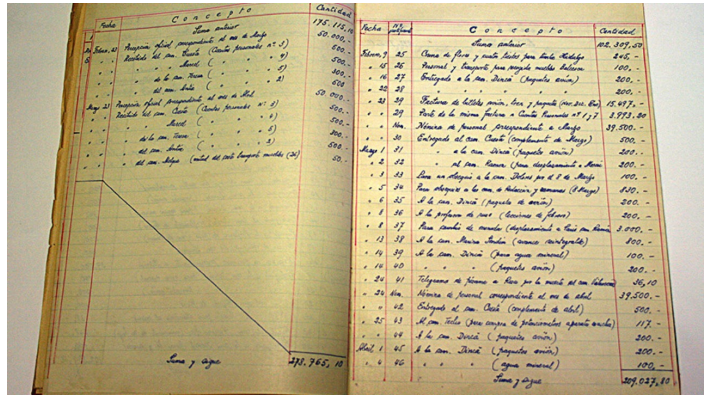
Avances

- Producción de datos.
- Capacidad de almacenamiento.

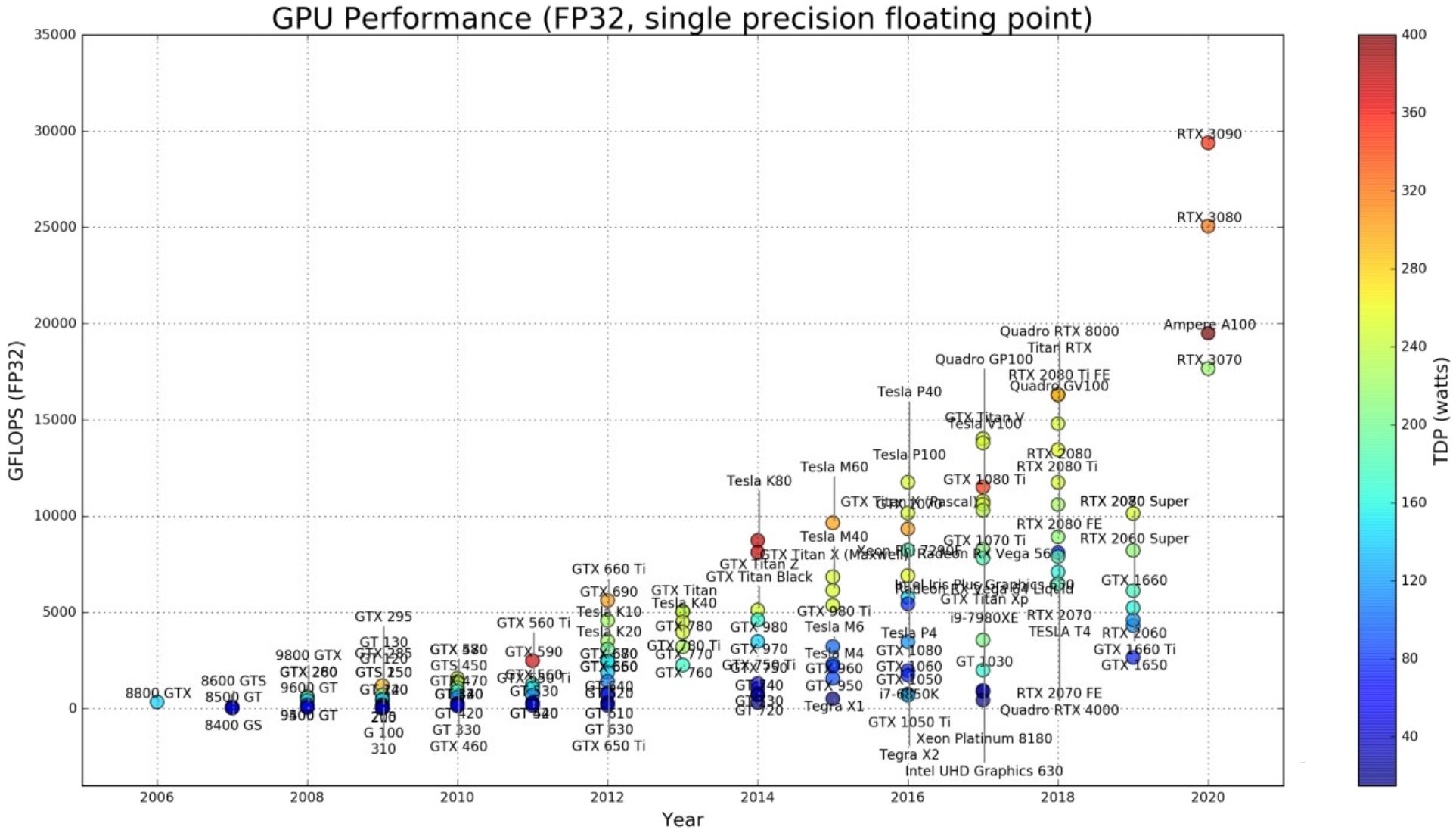


Avances

- Producción de datos.
- Capacidad de almacenamiento.
- Capacidad de procesamiento.



Comparativa GPUs

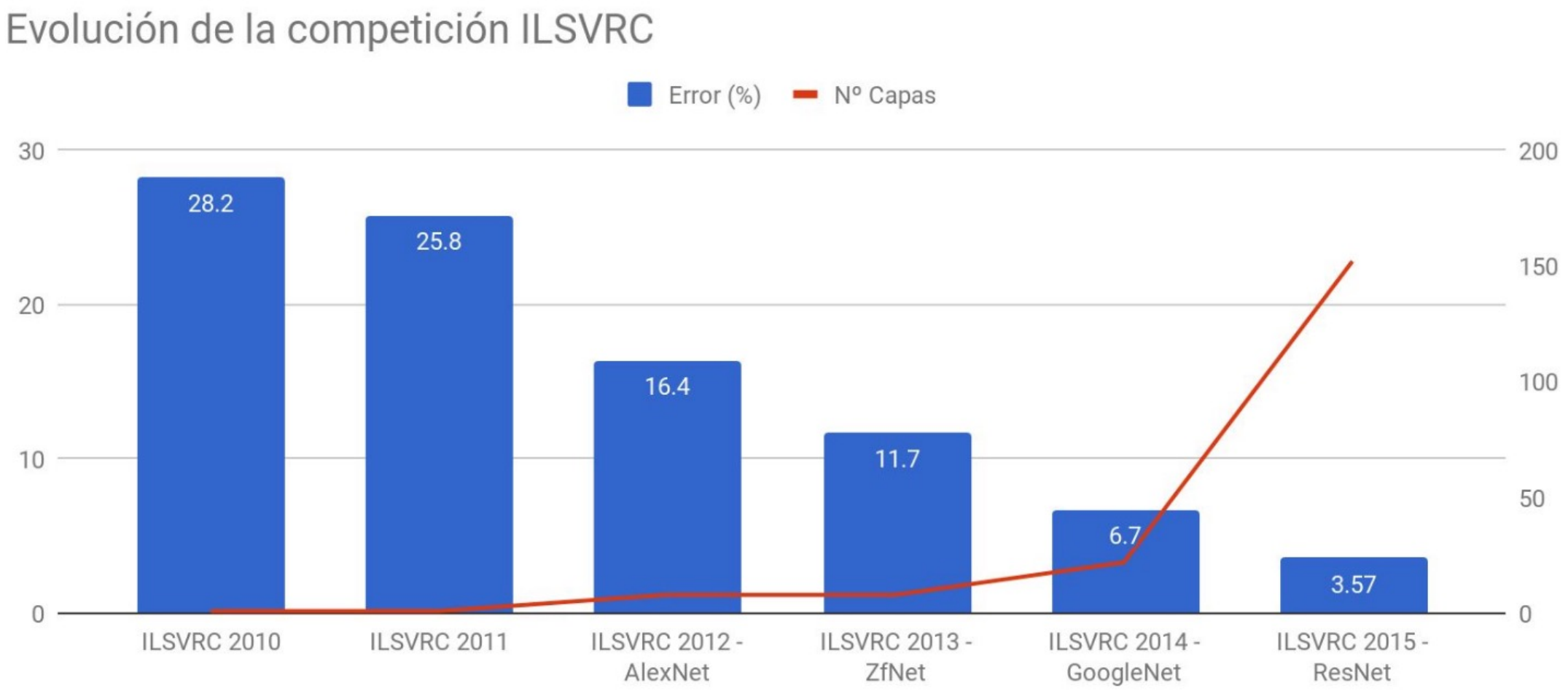


Fuente: https://blog.inten.to/hardware-for-deep-learning-part-3-gpu-8906c1644664

Datos

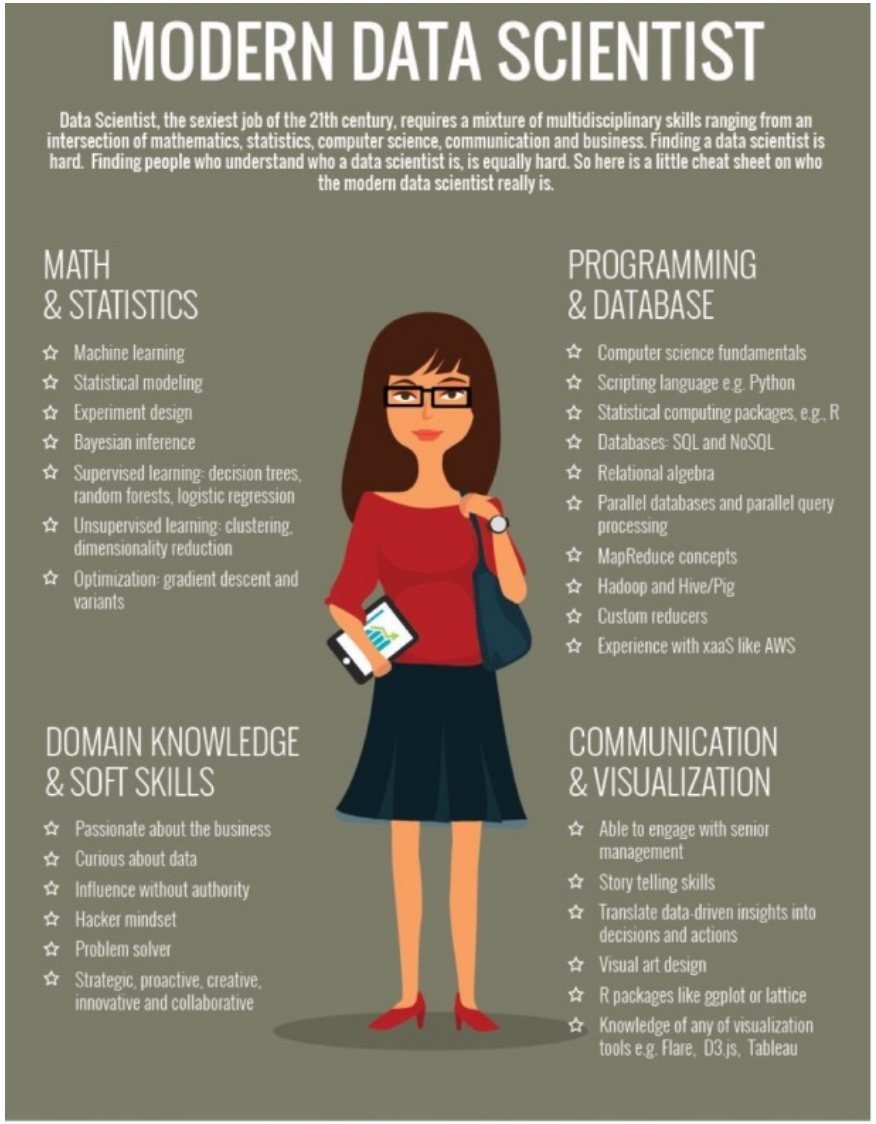

LOS DATOS SON
EL NUEVO PETRÓLEO
• ADQUIEREN un valor estratégico para las empresas
• UNA NUEVA clase de activo

Evolución de los modelos



The sexiest job of the century!!

Harvard Business Review

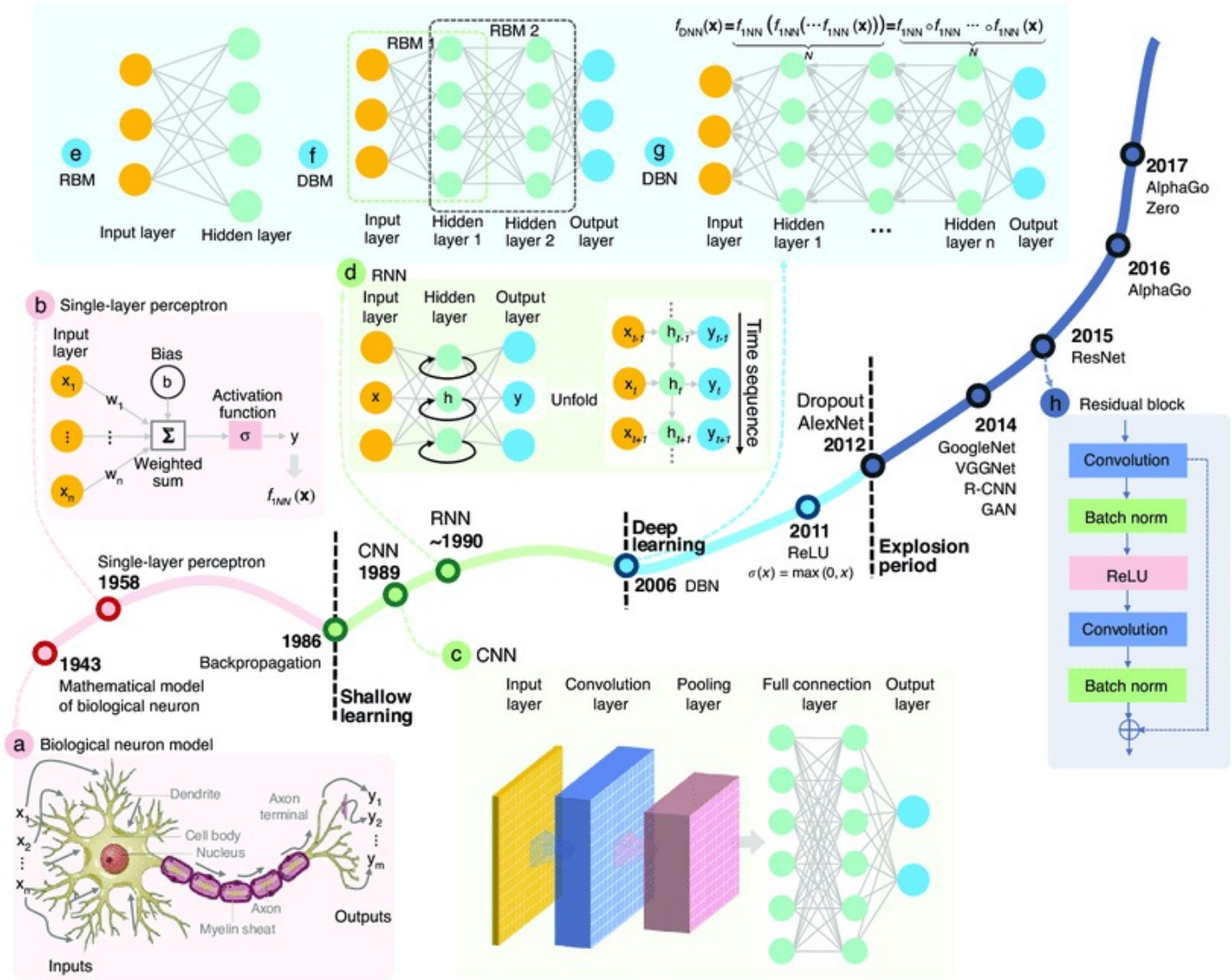


Deep Learning

ETS de Ingeniería Informática

UNED

Evolución de DNN



Generación de texto

at first:

```
tyntd-iafhatawiaoihrdemot  lytdws  e ,tfti, astai f ogoh eoase rrranbyne 'nhthnee e
plia tklrgd t o idoe ns,smtt   h ne etie h,hregtrs nigtike,aoaenns lng
```

train more

```
"Tmont thithey" fomesscerliund
Keushey. Thom here
sheulke, anmerenith ol sivh I lalterthend Bleipile shuwy fil on aseterlome
coaniogennc Phe lism thond hon at. MeiDimorotion in ther thize."
```

train more

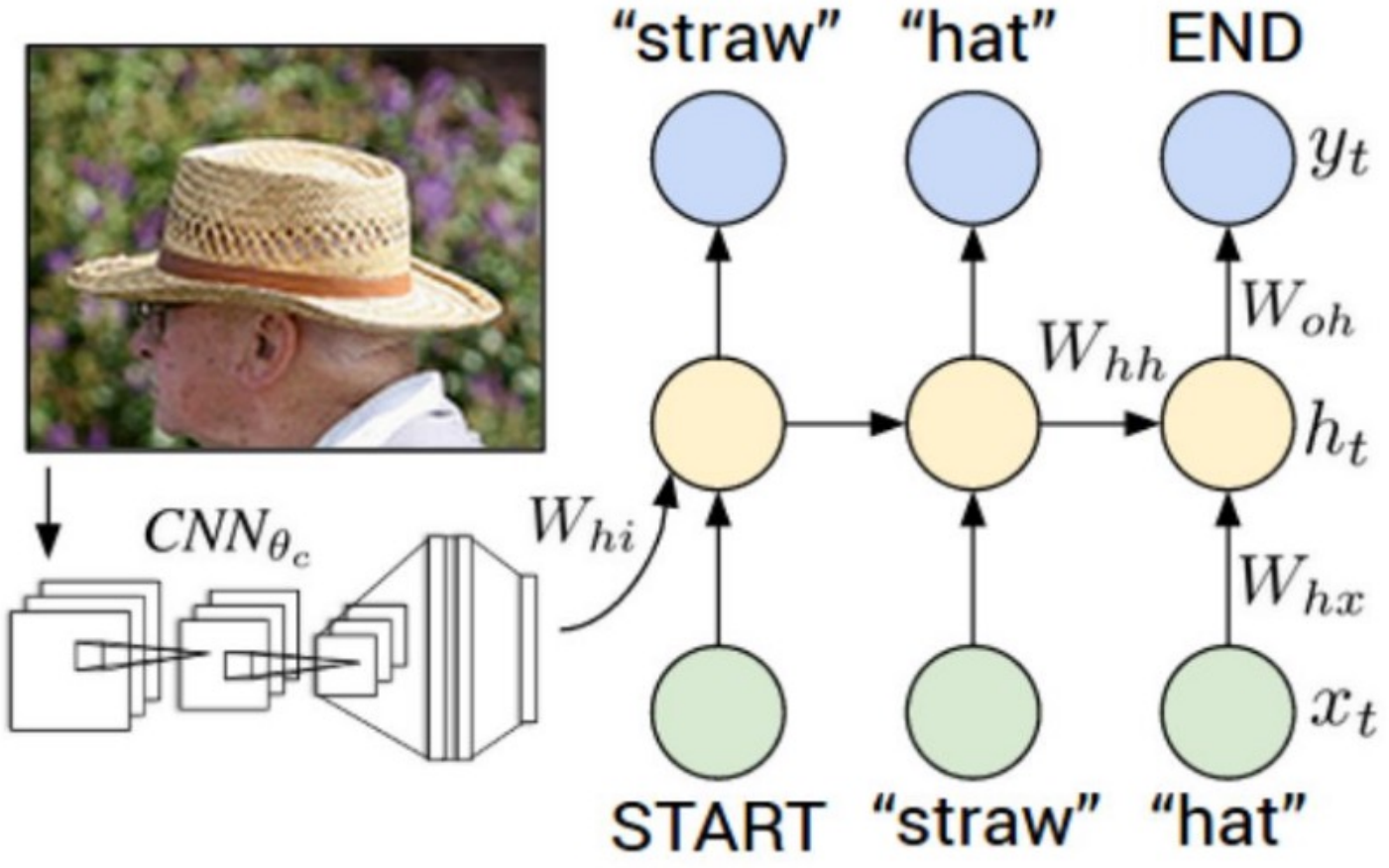
```
Aftair fall unsuch that the hall for Prince Velzonski's that me of
her hearly, and behs to so arwage fiving were to it beloge, pavu say falling misfort
how, and Gogition is so overelical and ofter.
```

train more

```
"Why do what that day," replied Natasha, and wishing to himself the fact the
princess, Princess Mary was easier, fed in had oftened him.
Pierre aking his soul came to the packs and drove up his father-in-law women.
```

Fuente: http://cs231n.stanford.edu/index.html

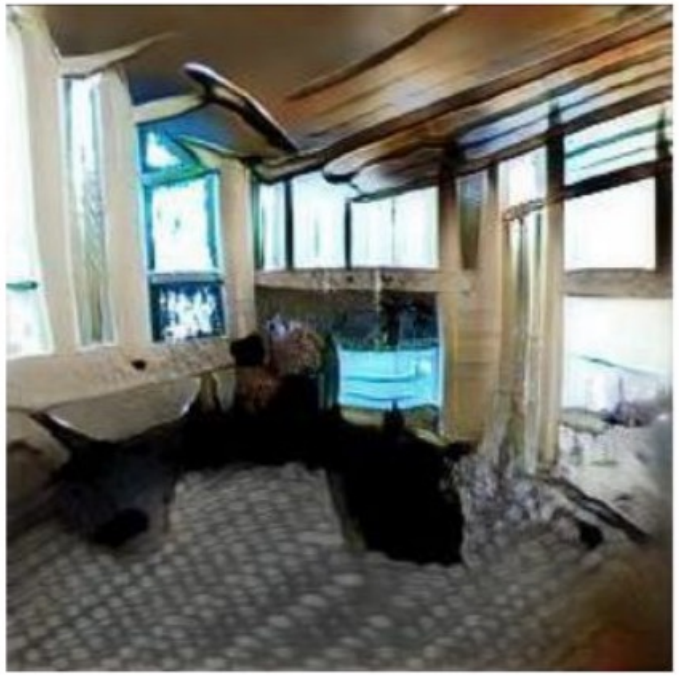
Captar imagen



Fuente: Karpathy, A., & Fei-Fei, L. (2015). Deep visual-semantic alignments for generating image descriptions. In Proceedings of the IEEE conference on computer vision and pattern recognition.

Síntesis de imágenes

- Salas amuebladas con suelos de madera



Zhang, H., Xu, T., Li, H., Zhang, S., Huang, X., Wang, X., & Metaxas, D. (2017, October). *Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks*. In IEEE Int. Conf. Comput. Vision (ICCV) (pp. 5907-5915).

Síntesis de imágenes

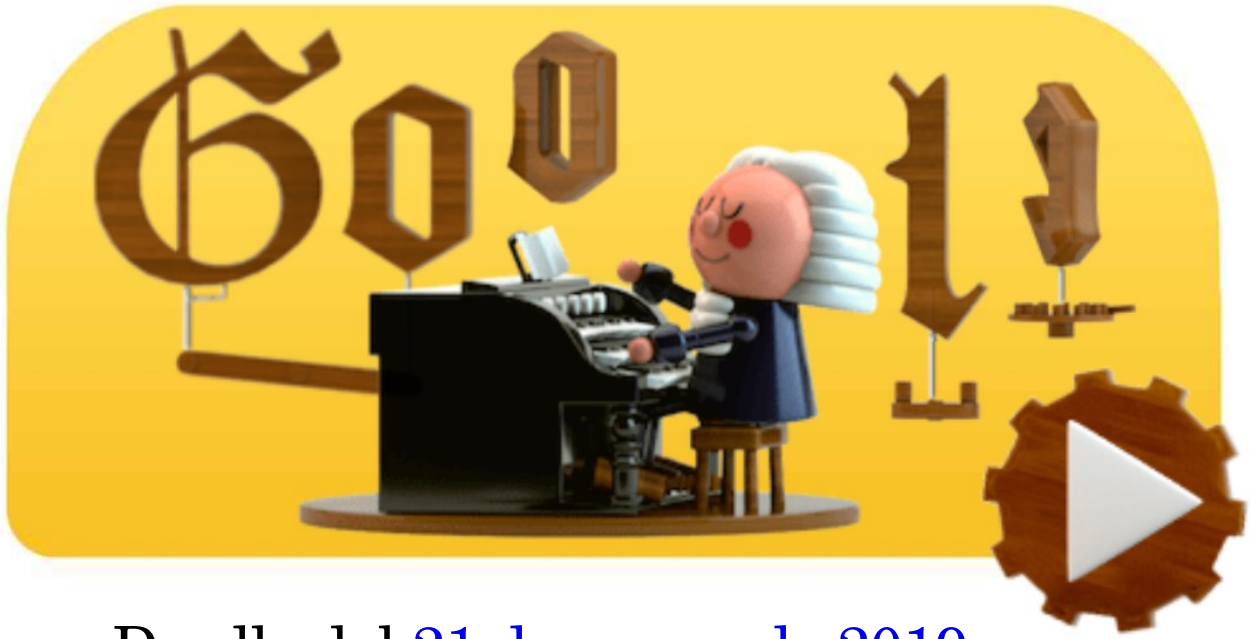
- Un pequeño pájaro con cabeza roja cuyas alas van desde el gris en la raíz hasta el rojo de las puntas

Zhang, H., Xu, T., Li, H., Zhang, S., Huang, X., Wang, X., & Metaxas, D. (2017, October). *Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks*. In IEEE Int. Conf. Comput. Vision (ICCV) (pp. 5907-5915).

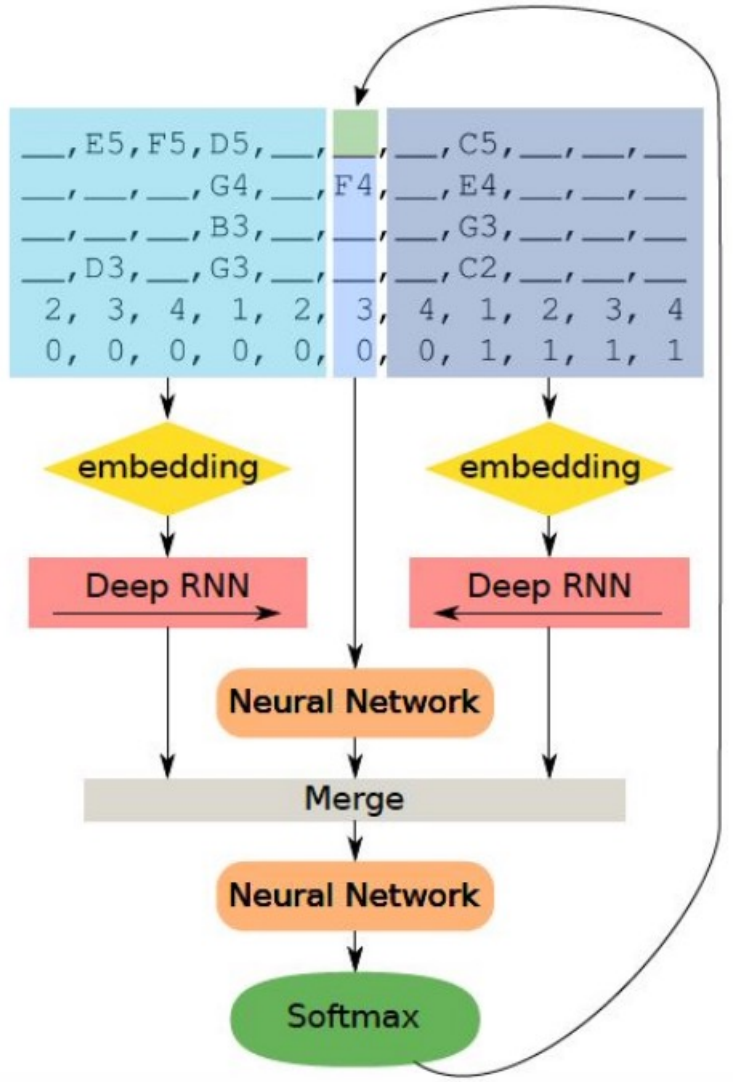
Imágenes de caras sintéticas

Karras, T., Laine, S., & Aila, T. (2018). *A style-based generator architecture for generative adversarial networks*. arXiv preprint arXiv:1812.04948.

DeepBach



Doodle del 21 de marzo de 2019
¡Funcionaba con CNNs!

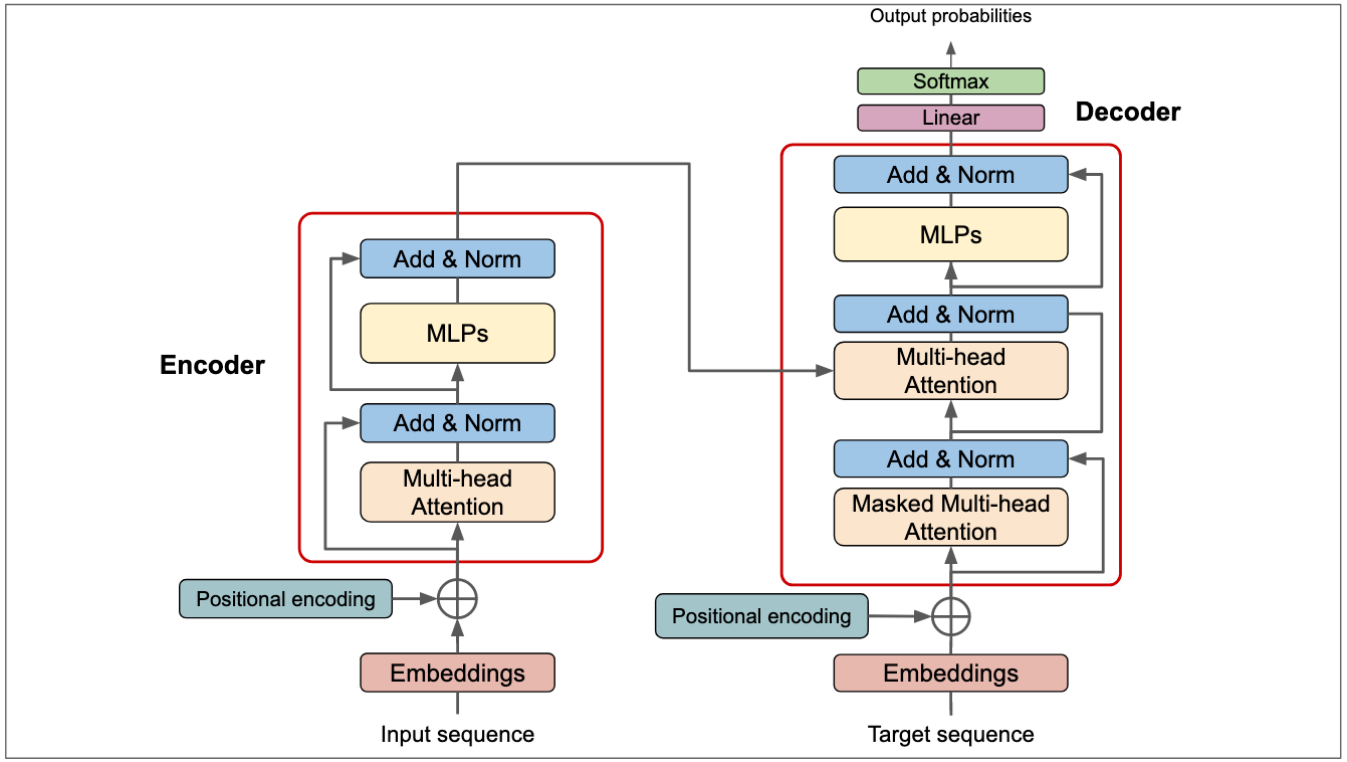


Fuente: Hadjeres, G., Pachet, F., & Nielsen, F. (2017). DeepBach: a steerable model for Bach chorales generation. In Proceedings of the 34th International Conference on Machine Learning

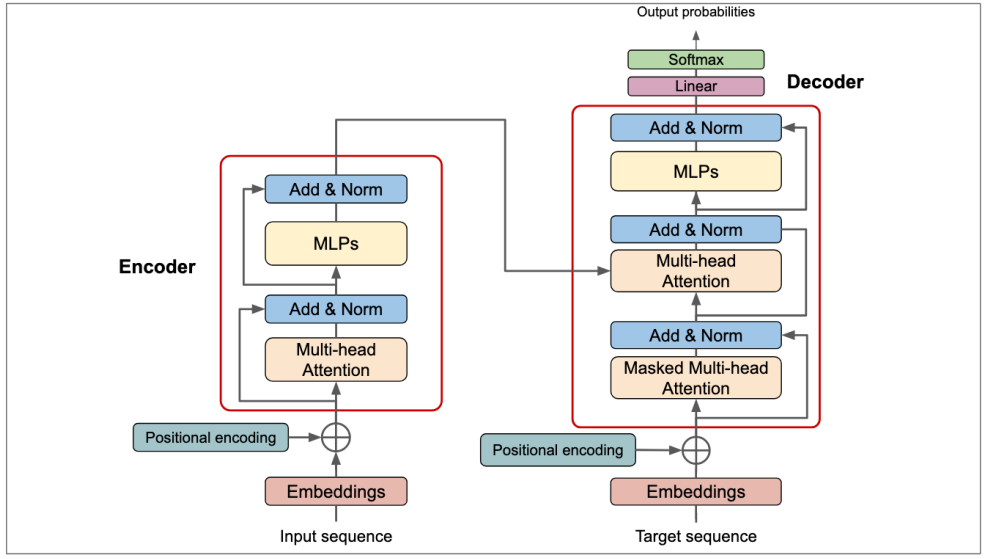
¡Como pasa el tiempo!

¡Como pasa el tiempo!

La guerra de los LLM



La guerra de los LLM



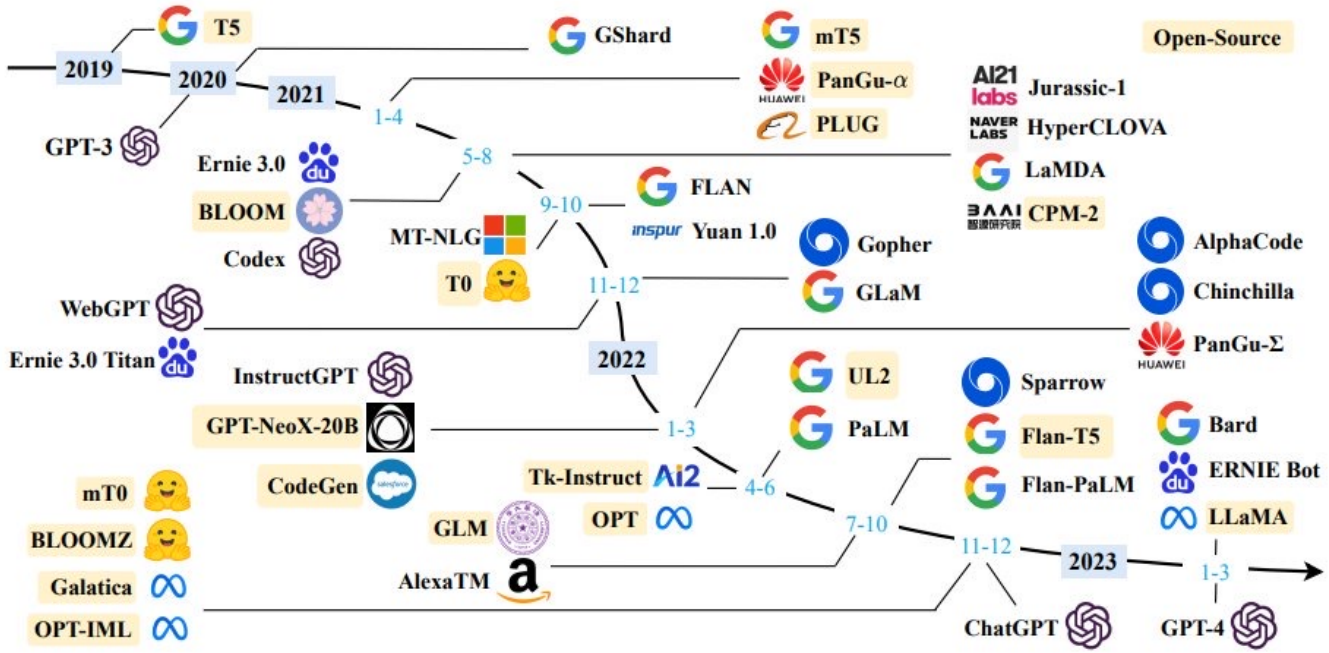
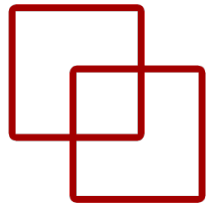


Fig. 1. A timeline of existing large language models (having a size larger than 10B) in recent years. We mark the open-source LLMs in yellow color.

Pero volvamos al inicio

Regression vs. Classification



	A	B	C	D	E	F	G	H	I
1	preg	plas	pres	skin	insu	mass	pedi	age	class
2	1	85	66	29	0	26.6	351	31	tested_negative
3	5	116	74	0	0	25.6	201	30	tested_negative
4	10	115	0	0	0	35.3	134	29	tested_negative
5	4	110	92	0	0	37.6	191	30	tested_negative
6	10	139	80	0	0	27.1	1441	57	tested_negative
7	8	99	84	0	0	35.4	388	50	tested_negative
8	5	117	92	0	0	34.1	337	38	tested_negative
9	5	109	75	26	0	36	546	60	tested_negative

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PTRATIO	B	LSTAT	class
2	0.00632	18	2.31	0	538	6575	65.2	4.09	1	296	15.3	396.9	4.98	24
3	0.02731	0	7.07	0	469	6421	78.9	4.9671	2	242	17.8	396.9	9.14	21.6
4	0.02729	0	7.07	0	469	7185	61.1	4.9671	2	242	17.8	392.83	4.03	34.7
5	0.03237	0	2.18	0	458	6998	45.8	6.0622	3	222	18.7	394.63	2.94	33.4
6	0.06905	0	2.18	0	458	7147	54.2	6.0622	3	222	18.7	396.9	5.33	36.2
7	0.02985	0	2.18	0	458	6.43	58.7	6.0622	3	222	18.7	394.12	5.21	28.7

Regression vs. Classification

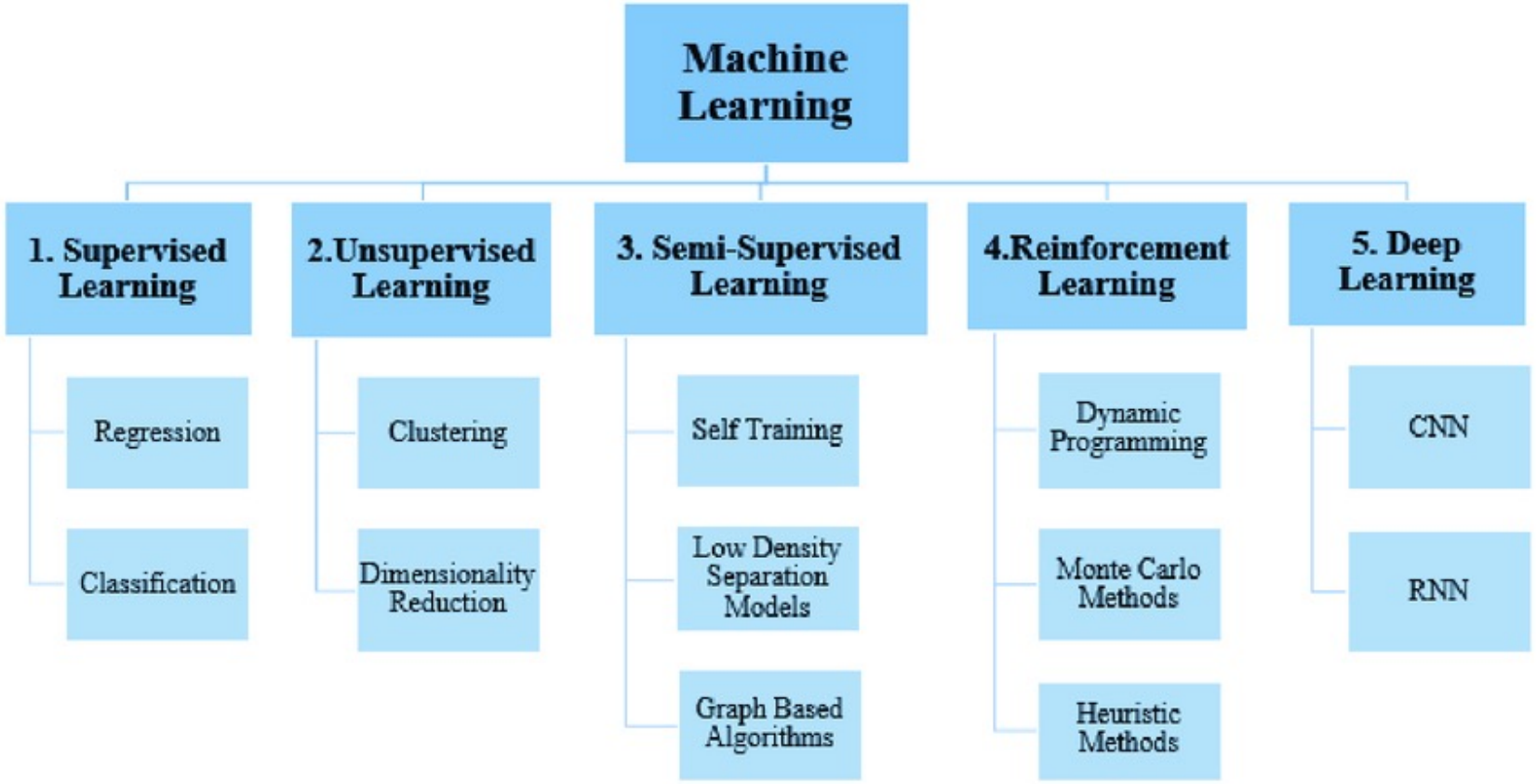
	A	B	C	D	E	F	G	H	I
1	preg	plas	pres	skin	insu	mass	pedi	age	class
2	1	85	66	29	0	26.6	351	31	tested_negative
3	5	116	74	0	0	25.6	201	30	tested_negative
4	10	115	0	0	0	35.3	134	29	tested_negative
5	4	110	92	0	0	37.6	191	30	tested_negative
6	10	139	80	0	0	27.1	1441	57	tested_negative
7	8	99	84	0	0	35.4	388	50	tested_negative
8	5	117	92	0	0	34.1	337	38	tested_negative
9	5	109	75	26	0	36	546	60	tested_negative

Classification

Regression

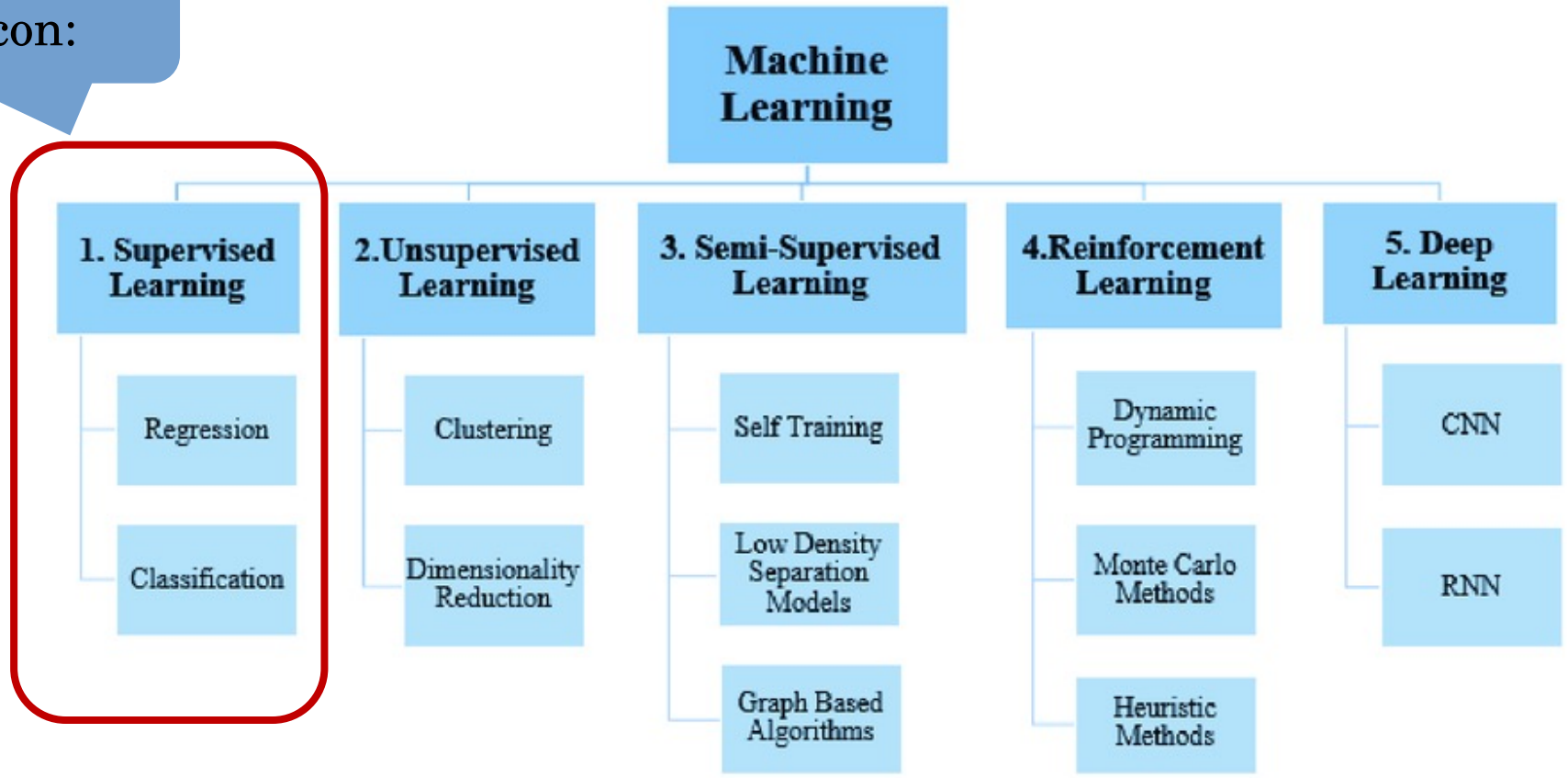
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PTRATIO	B	LSTAT	class
2	0.00632	18	2.31	0	538	6575	65.2	4.09	1	296	15.3	396.9	4.98	24
3	0.02731	0	7.07	0	469	6421	78.9	4.9671	2	242	17.8	396.9	9.14	21.6
4	0.02729	0	7.07	0	469	7185	61.1	4.9671	2	242	17.8	392.83	4.03	34.7
5	0.03237	0	2.18	0	458	6998	45.8	6.0622	3	222	18.7	394.63	2.94	33.4
6	0.06905	0	2.18	0	458	7147	54.2	6.0622	3	222	18.7	396.9	5.33	36.2
7	0.02985	0	2.18	0	458	6.43	58.7	6.0622	3	222	18.7	394.12	5.21	28.7

Machine Learning



Machine Learning

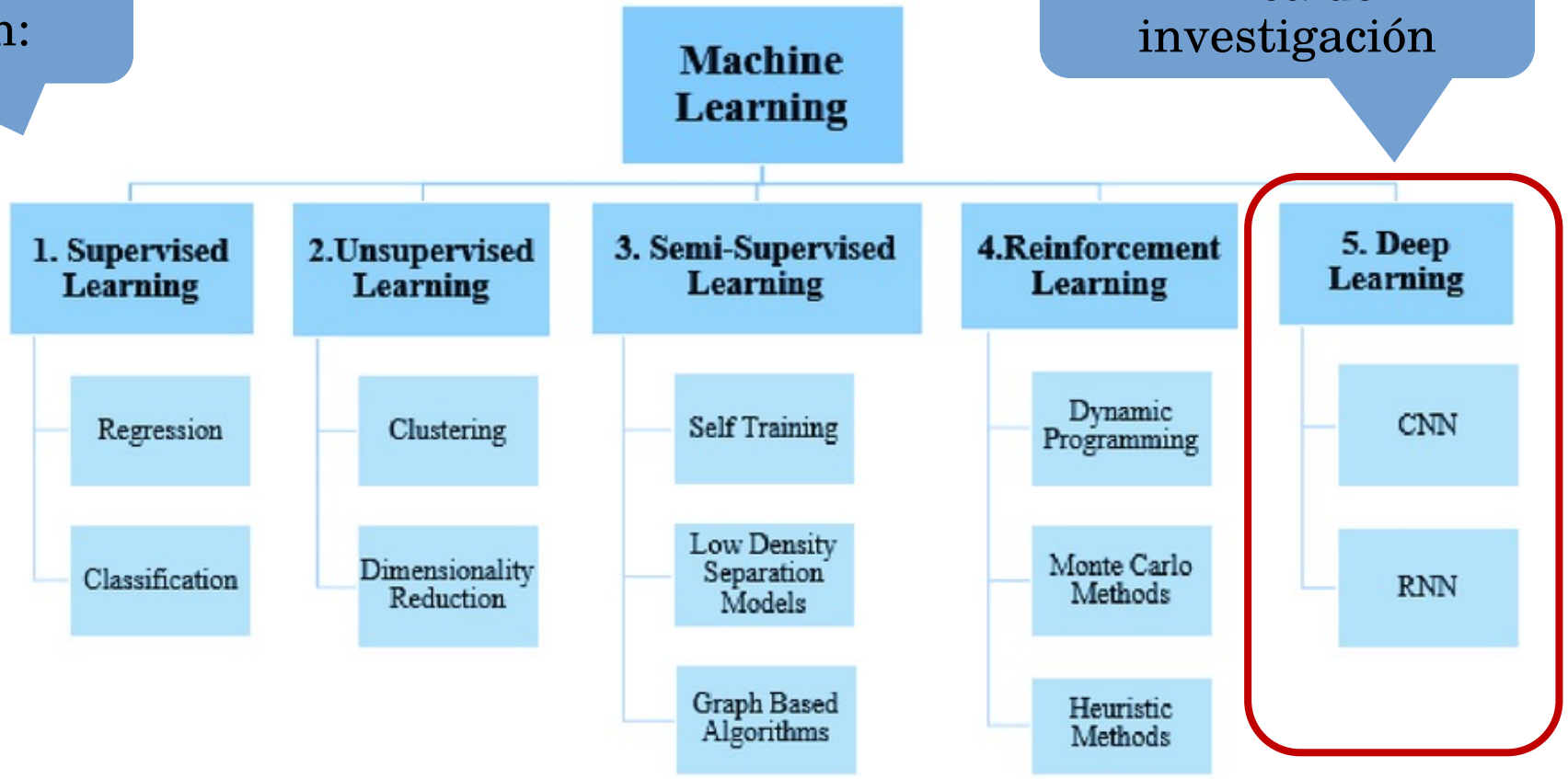
Normalmente se resuelven con:



Machine Learning

Normalmente se resuelven con:

Pero hay una nueva línea de investigación

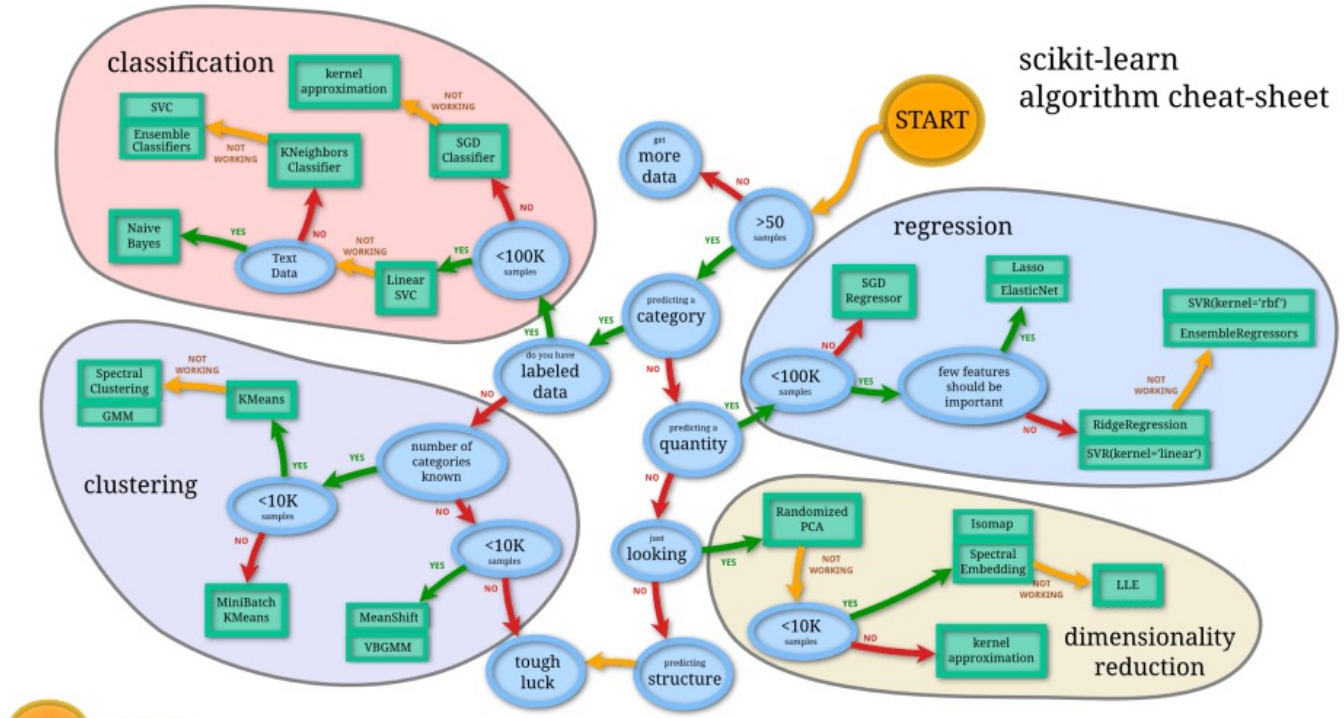


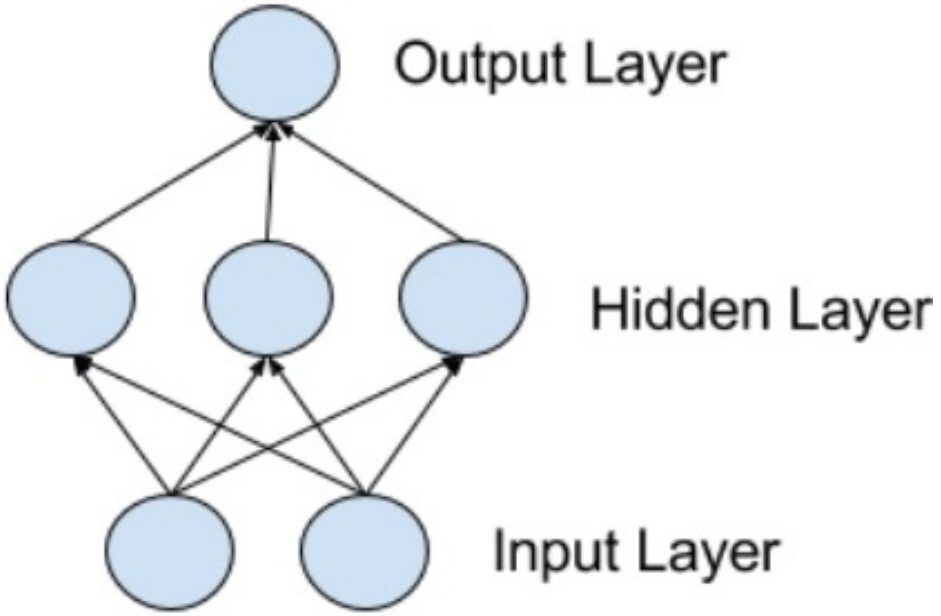
Arquitecturas neuronales

ETS de Ingeniería Informática

UNED

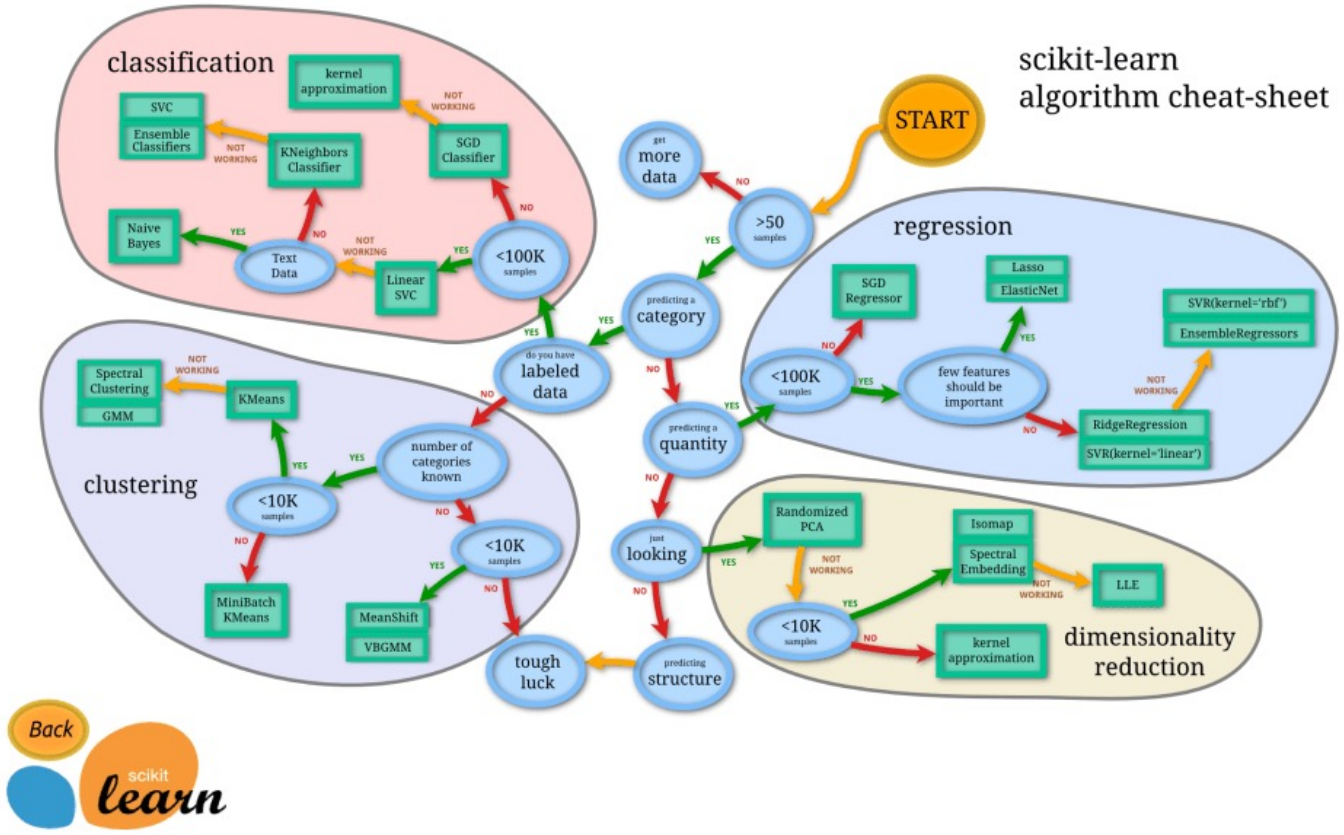
Algoritmos clásicos Vs. Multilayer Perceptrón

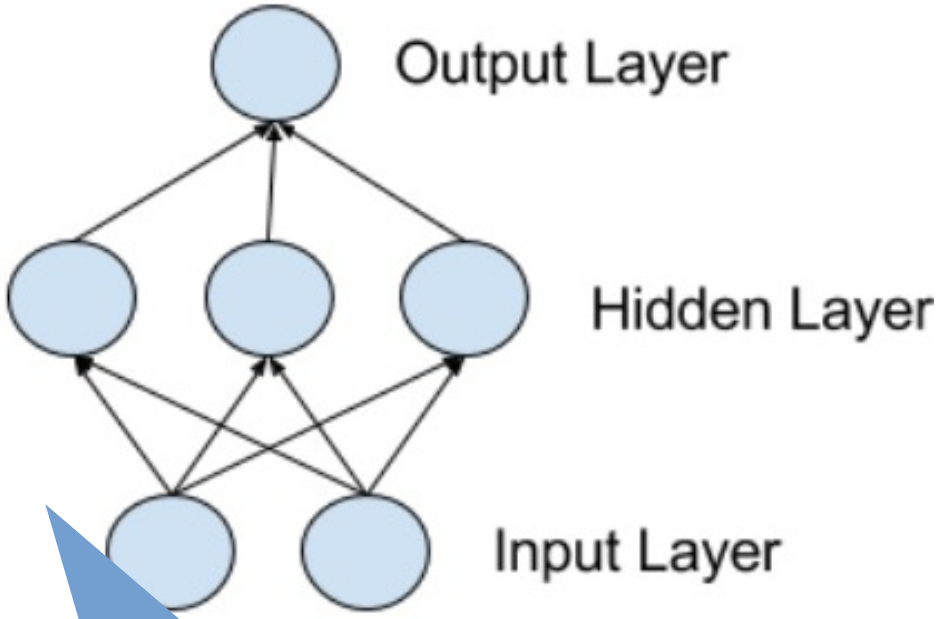




Back

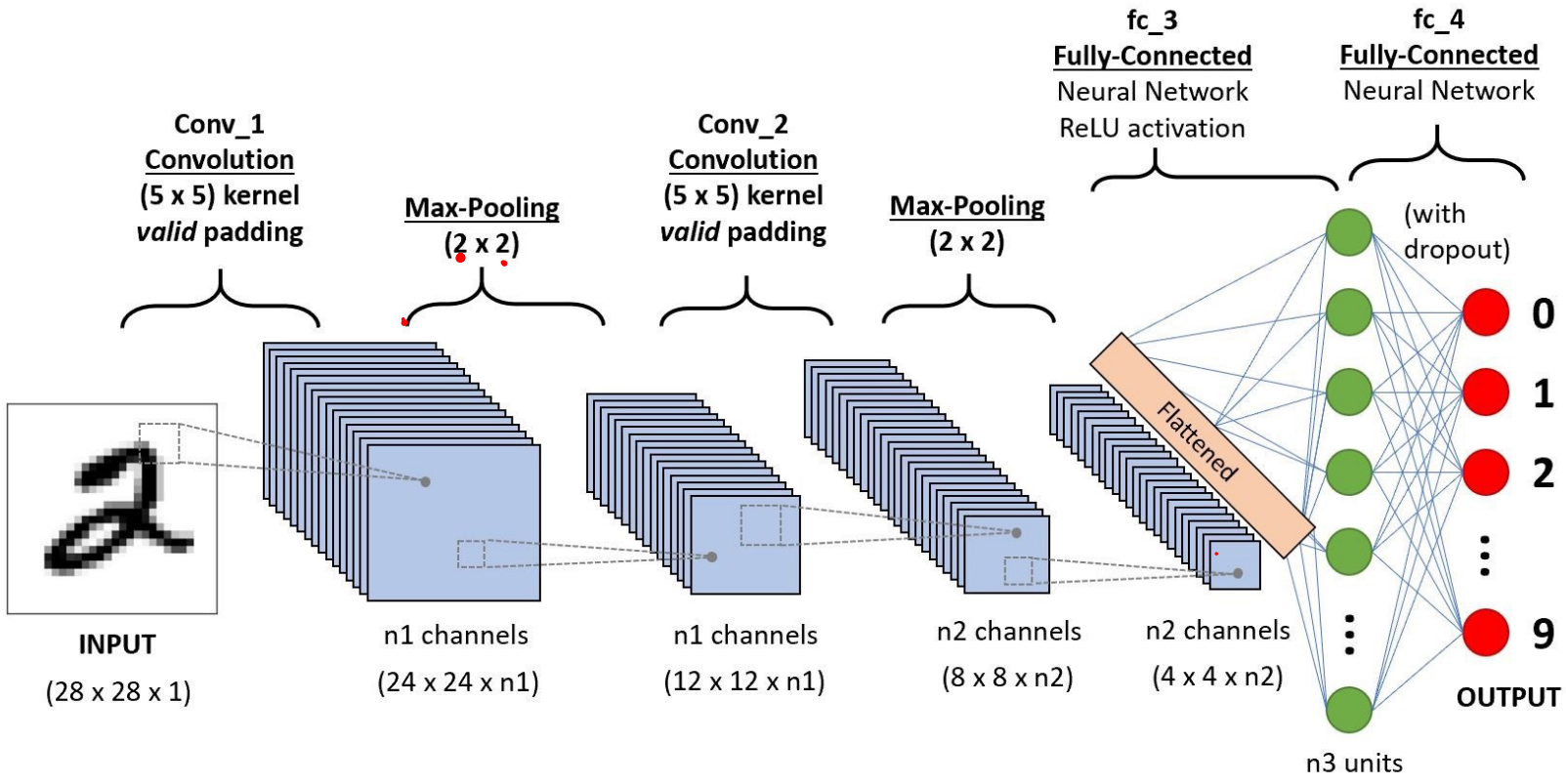
Algoritmos clásicos Vs. Multilayer Perceptrón





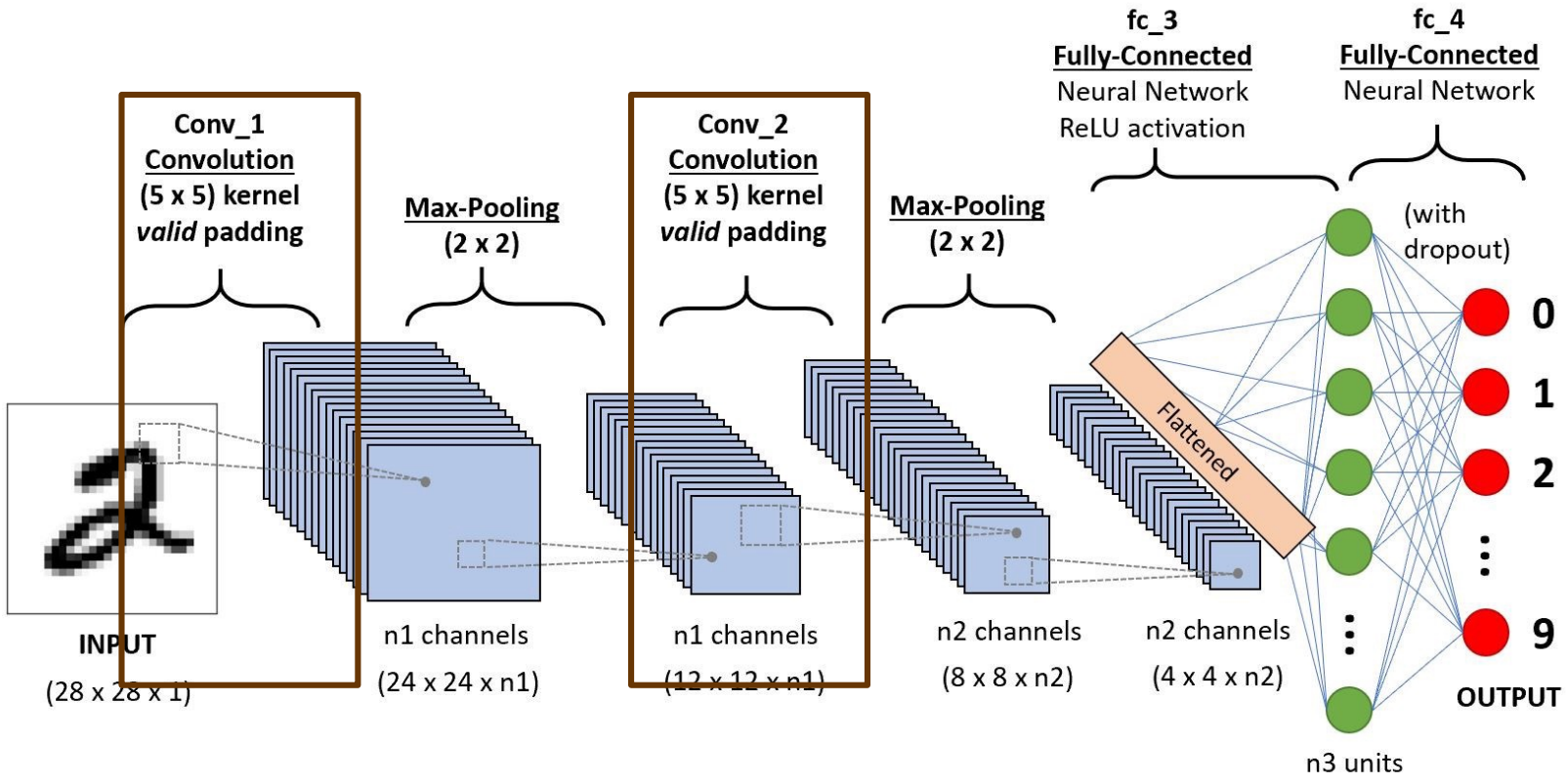
¡Para datos tabulares!

Redes Neuronales Convolucionales



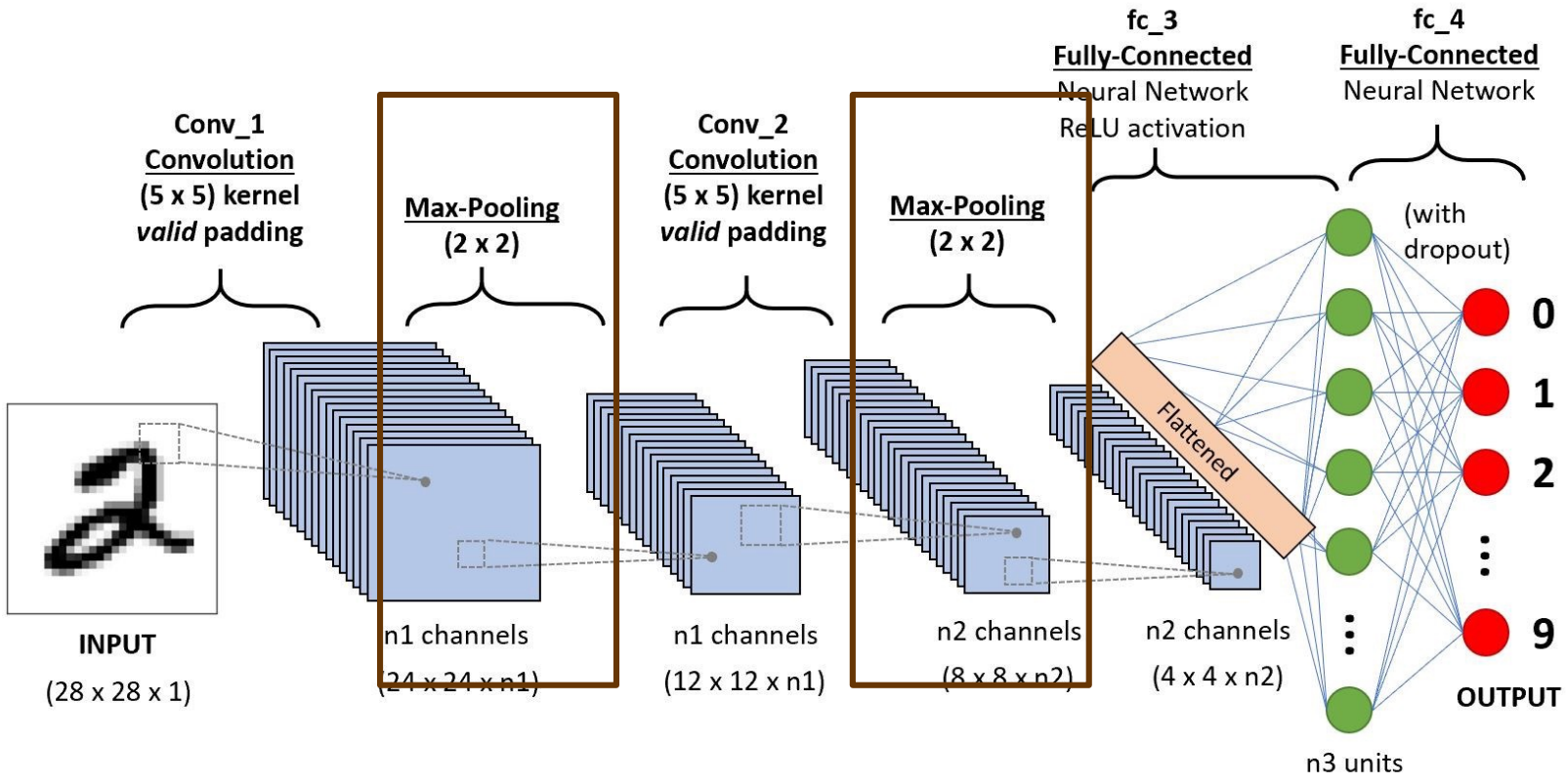
Redes Neuronales Convolucionales

Capas convolucionales

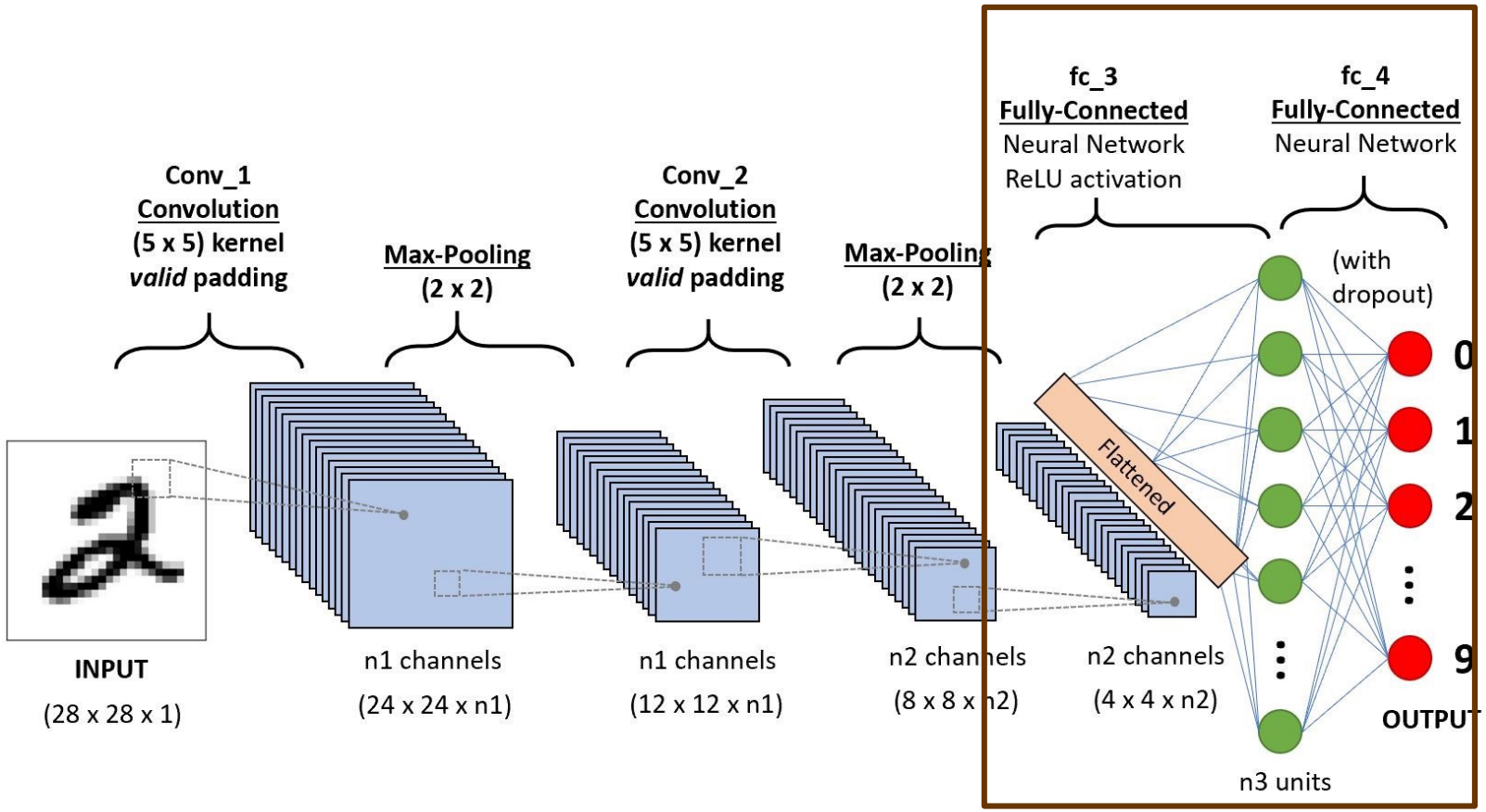


Redes Neuronales Convolucionales

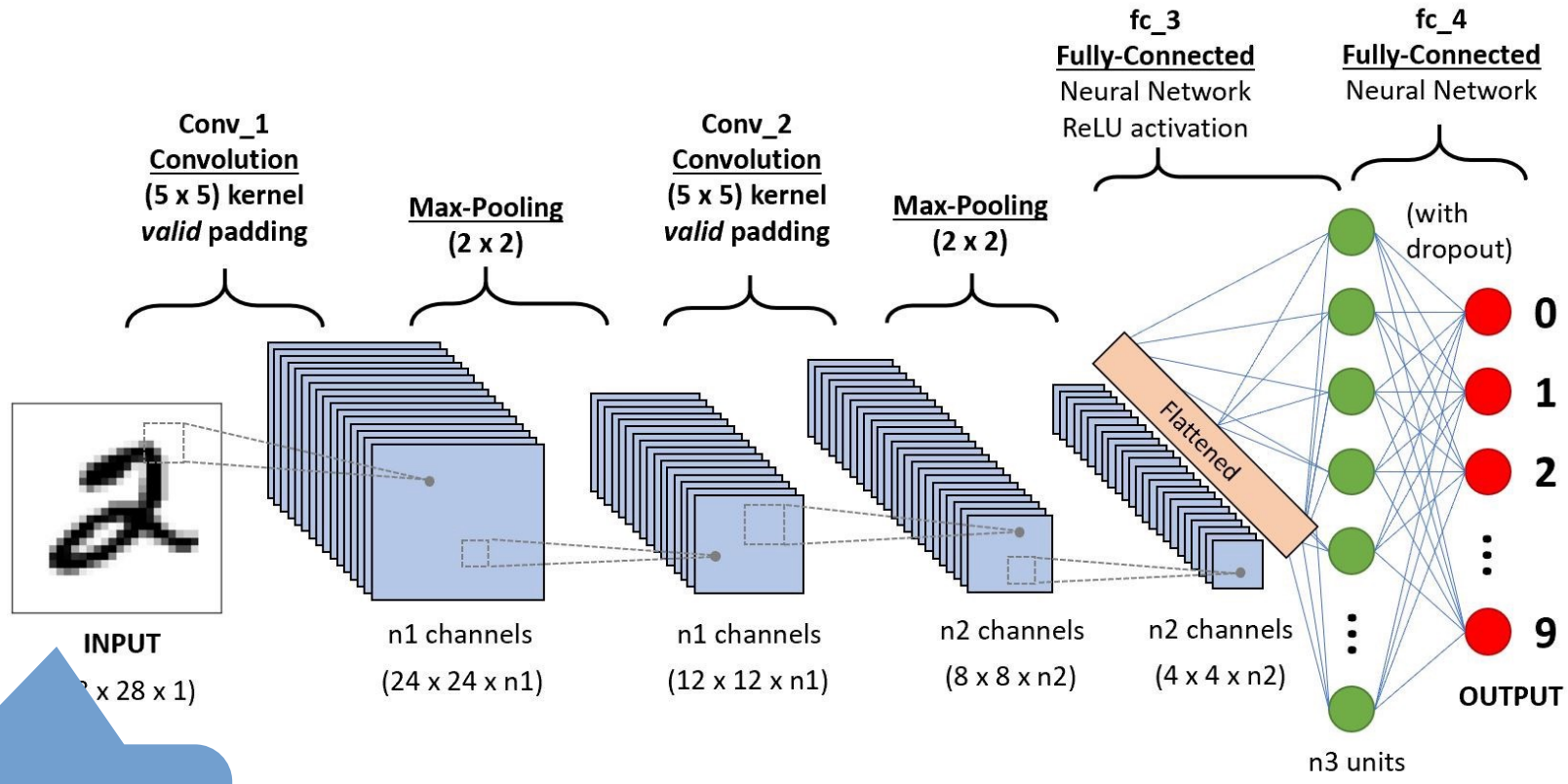
Capas de agrupación *(pooling)*



Redes Neuronales Convolucionales
Capas completamente conectadas



Redes Neuronales Convolucionales
Capas completamente conectadas



¡Para imágenes!

Vision Transformer (ViT)

¡Para imágenes!

Dosovitskiy, A., et al. (2020). An image is worth 16x16 words: Transformers for image recognition at scale.

Fundamentos de imágenes sintéticas

ETS de Ingeniería Informática

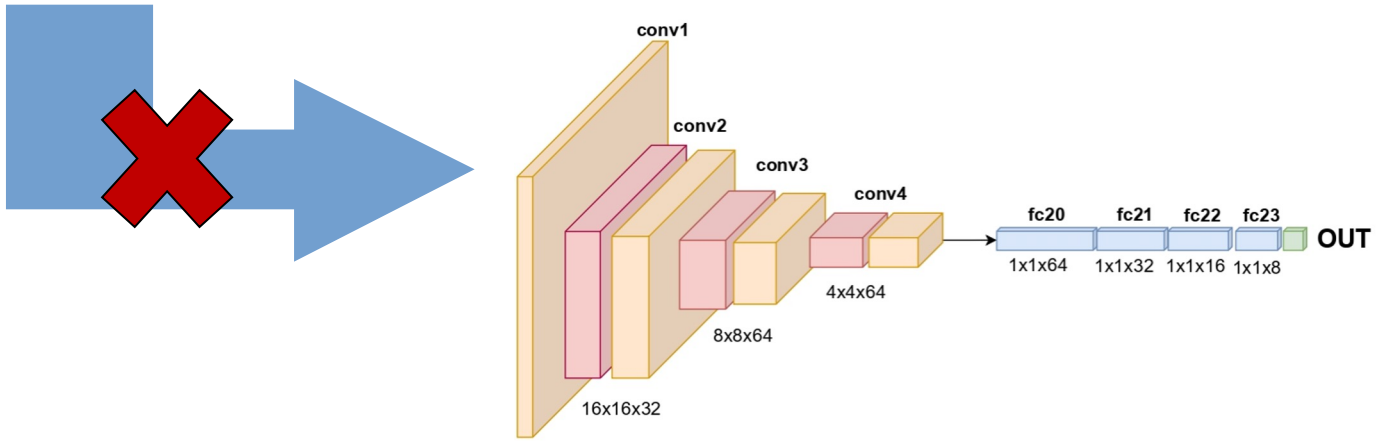
UNED

Transformación de datos

Be07	Be08	Be09	Be10	Be11	Sector
−65	−61	−74	−73	−67	1
−60	−57	−83	−62	−69	2
−66	−70	−78	−63	−73	3
...	...	...	...	...	...
−58	−66	−71	−73	−69	14
−60	−62	−73	−69	−57	15

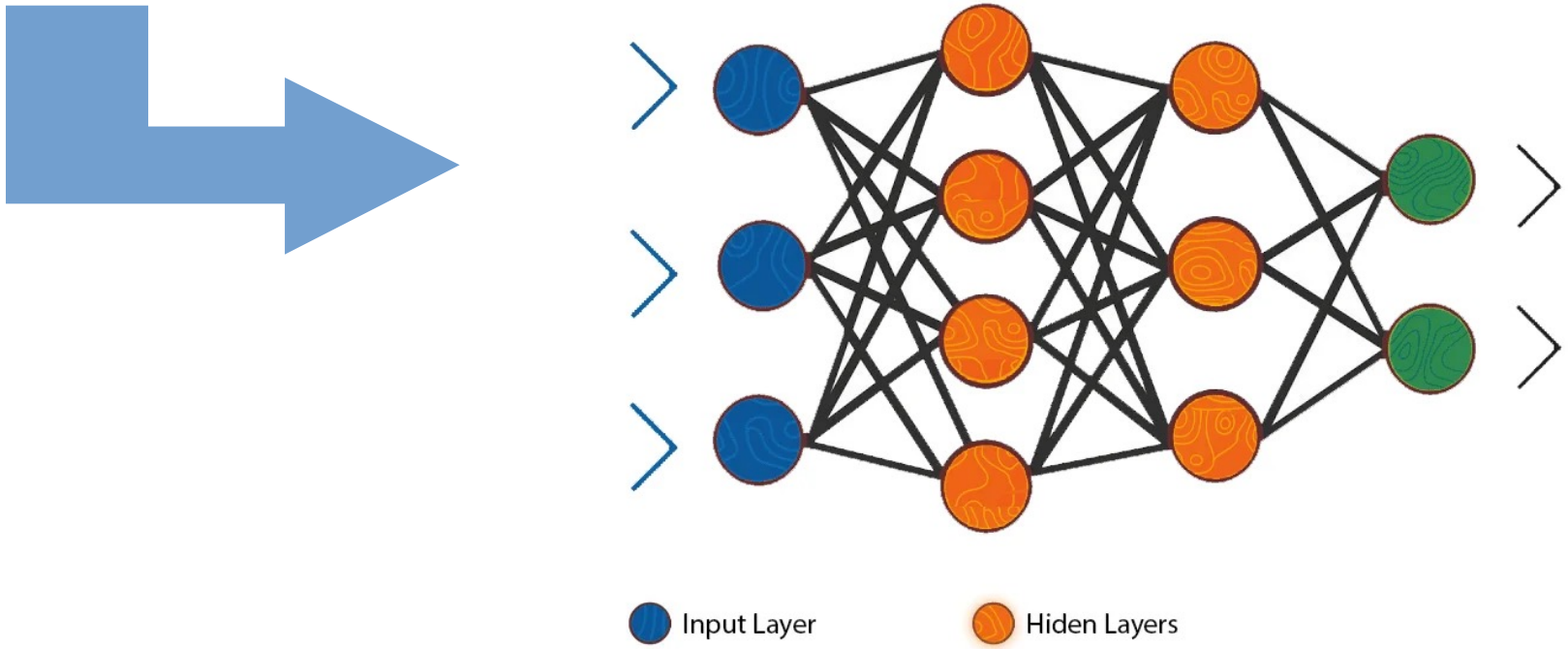
Transformación de datos

Be07	Be08	Be09	Be10	Be11	Sector
−65	−61	−74	−73	−67	1
−60	−57	−83	−62	−69	2
−66	−70	−78	−63	−73	3
...	...	...	...	...	...
−58	−66	−71	−73	−69	14
−60	−62	−73	−69	−57	15



Transformación de datos

Be07	Be08	Be09	Be10	Be11	Sector
−65	−61	−74	−73	−67	1
−60	−57	−83	−62	−69	2
−66	−70	−78	−63	−73	3
...	...	...	...	...	...
−58	−66	−71	−73	−69	14
−60	−62	−73	−69	−57	15

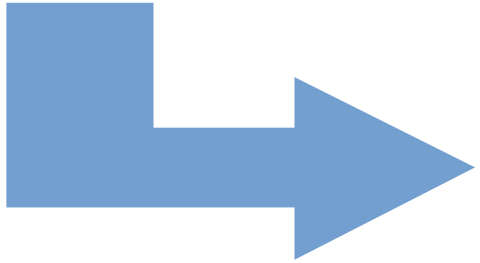


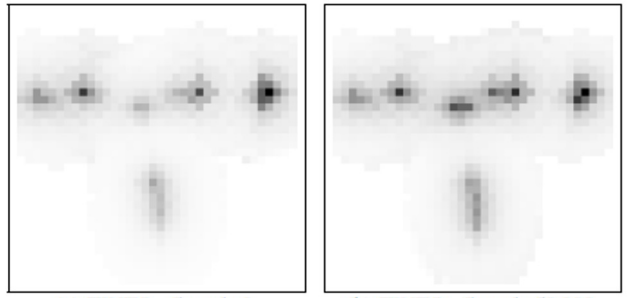
Transformación de datos

Be07	Be08	Be09	Be10	Be11	Sector
−65	−61	−74	−73	−67	1
−60	−57	−83	−62	−69	2
−66	−70	−78	−63	−73	3
...	...	...	...	...	...
−58	−66	−71	−73	−69	14
−60	−62	−73	−69	−57	15

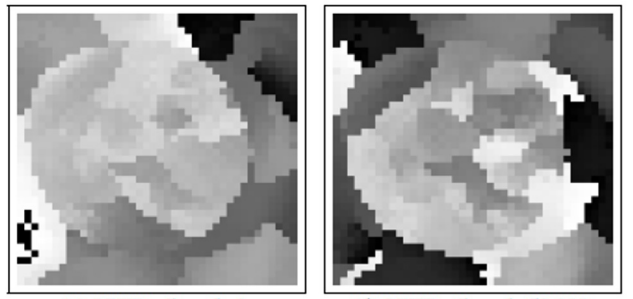
Transformación de datos

Be07	Be08	Be09	Be10	Be11	Sector
−65	−61	−74	−73	−67	1
−60	−57	−83	−62	−69	2
−66	−70	−78	−63	−73	3
...	...	...	...	...	...
−58	−66	−71	−73	−69	14
−60	−62	−73	−69	−57	15

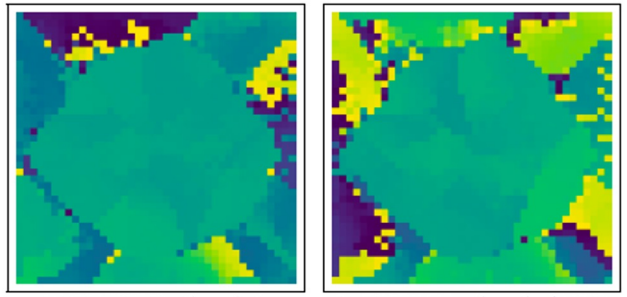




(a) TINTO - Sample 1. (b) TINTO - Sample 50,000.

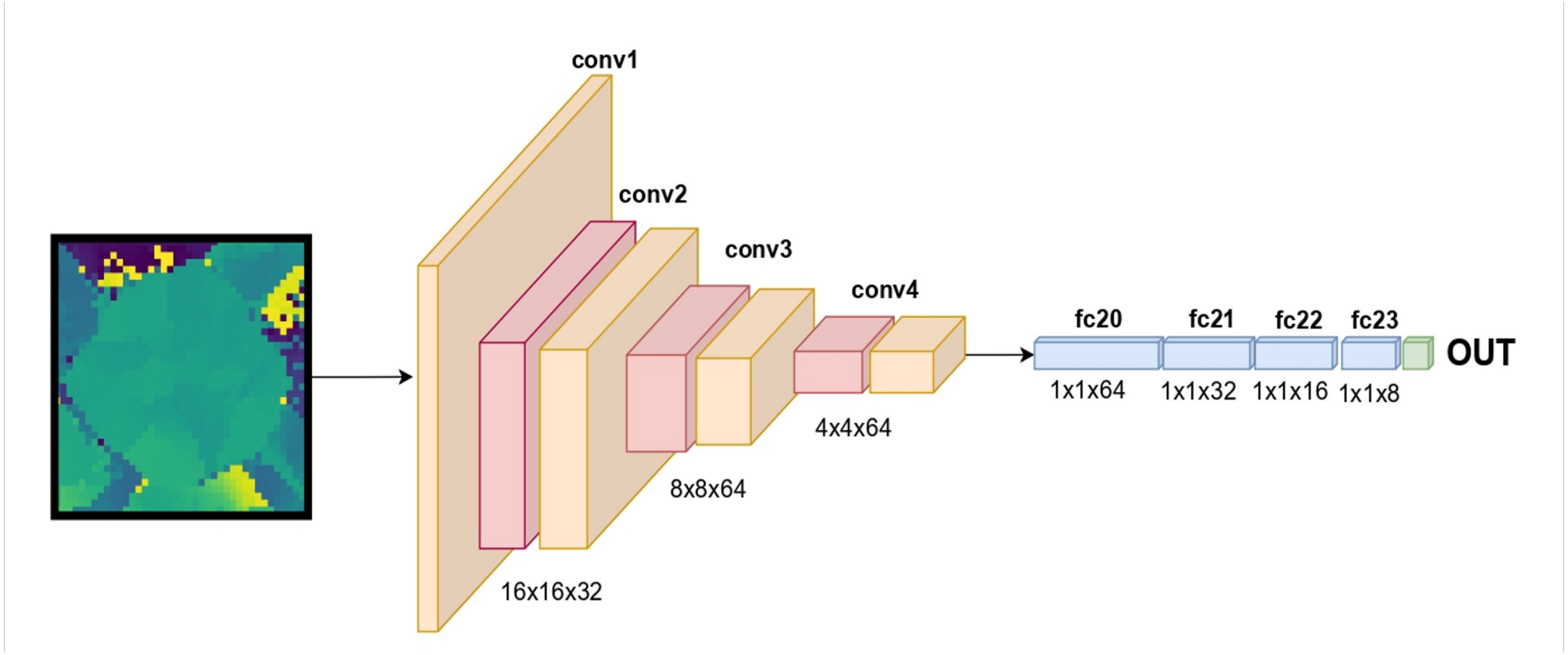


(c) IGTD - Sample 1. (d) IGTD - Sample 50,000.

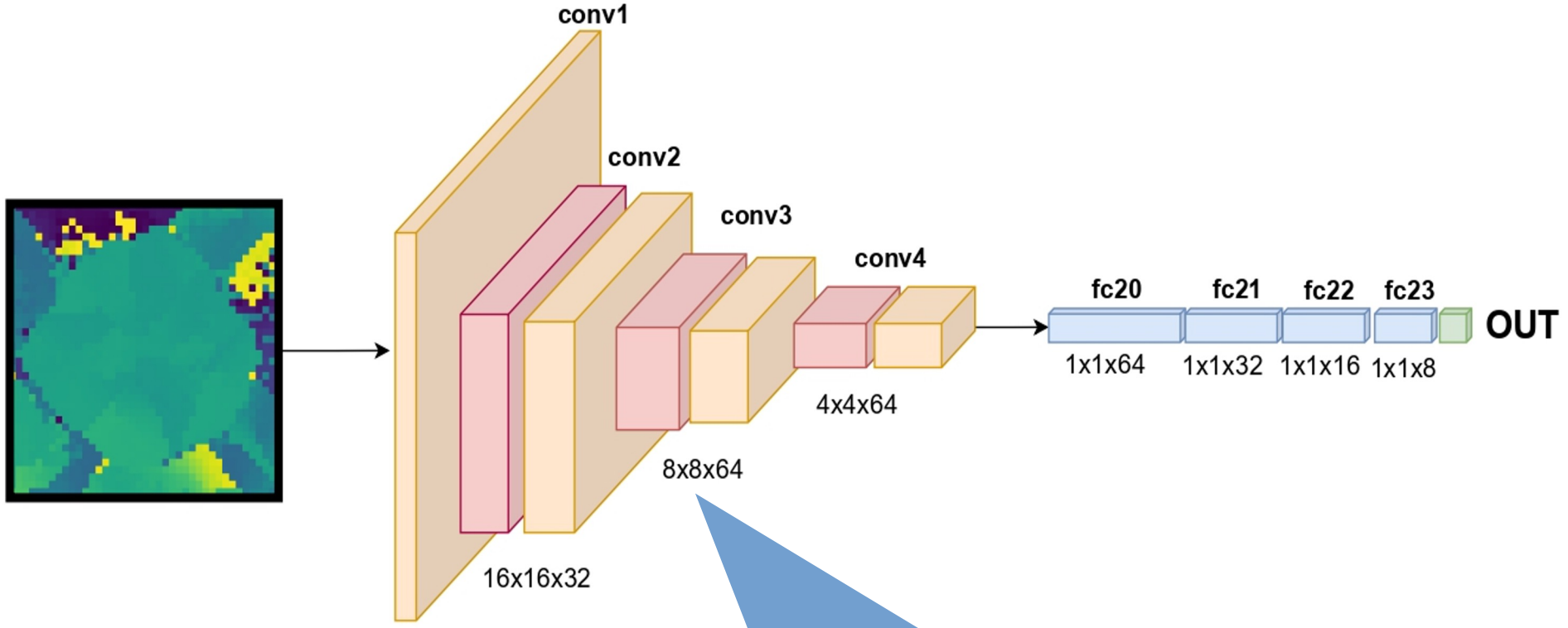


(e) REFINED - Sample 1. (f) REFINED - Sample 50,000.

CNN Puras

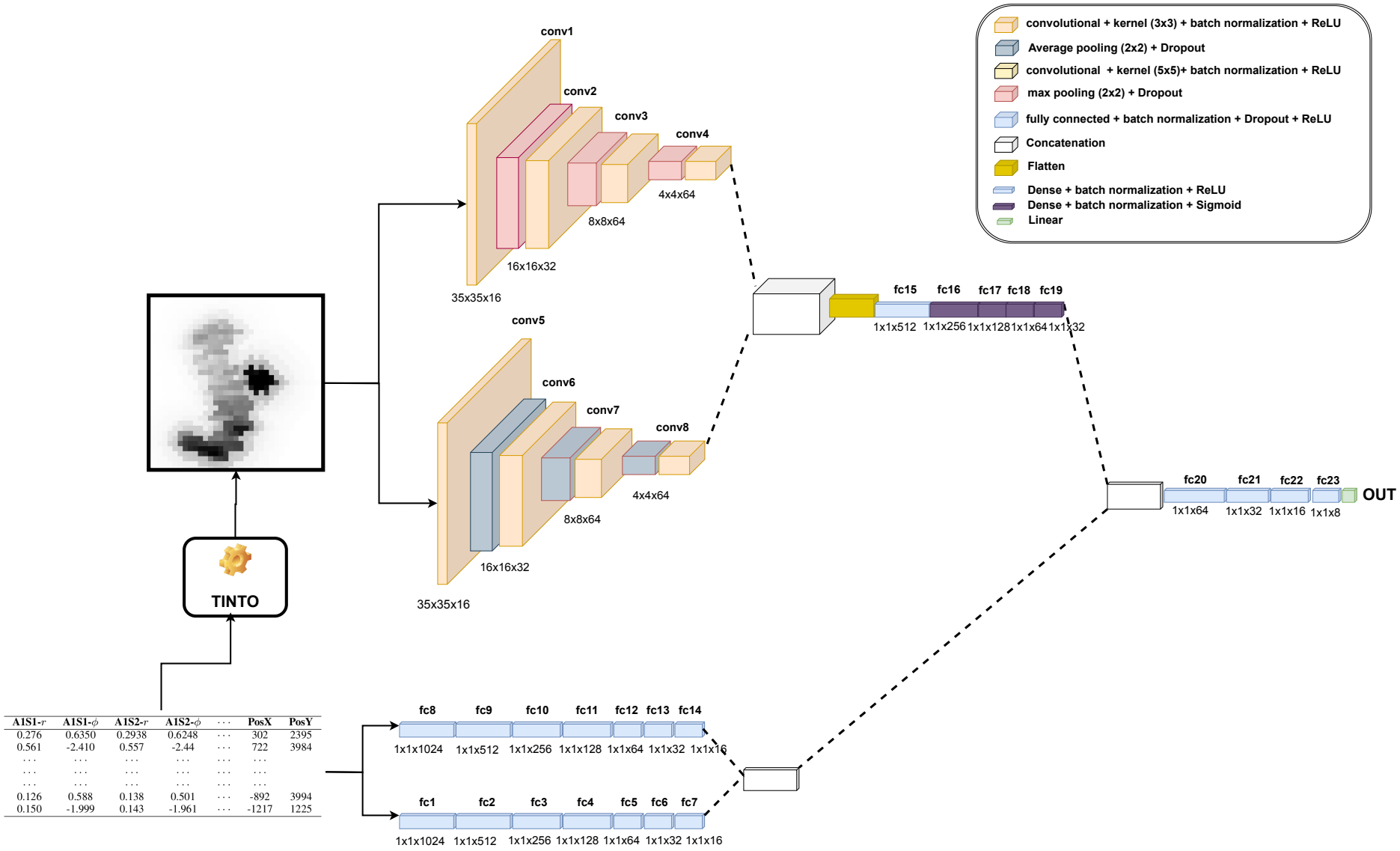

conv1
conv2
conv3
conv4
fc20
fc21
fc22
fc23
OUT
16x16x32
8x8x64
4x4x64
1x1x64
1x1x32
1x1x16
1x1x8

CNN Puras



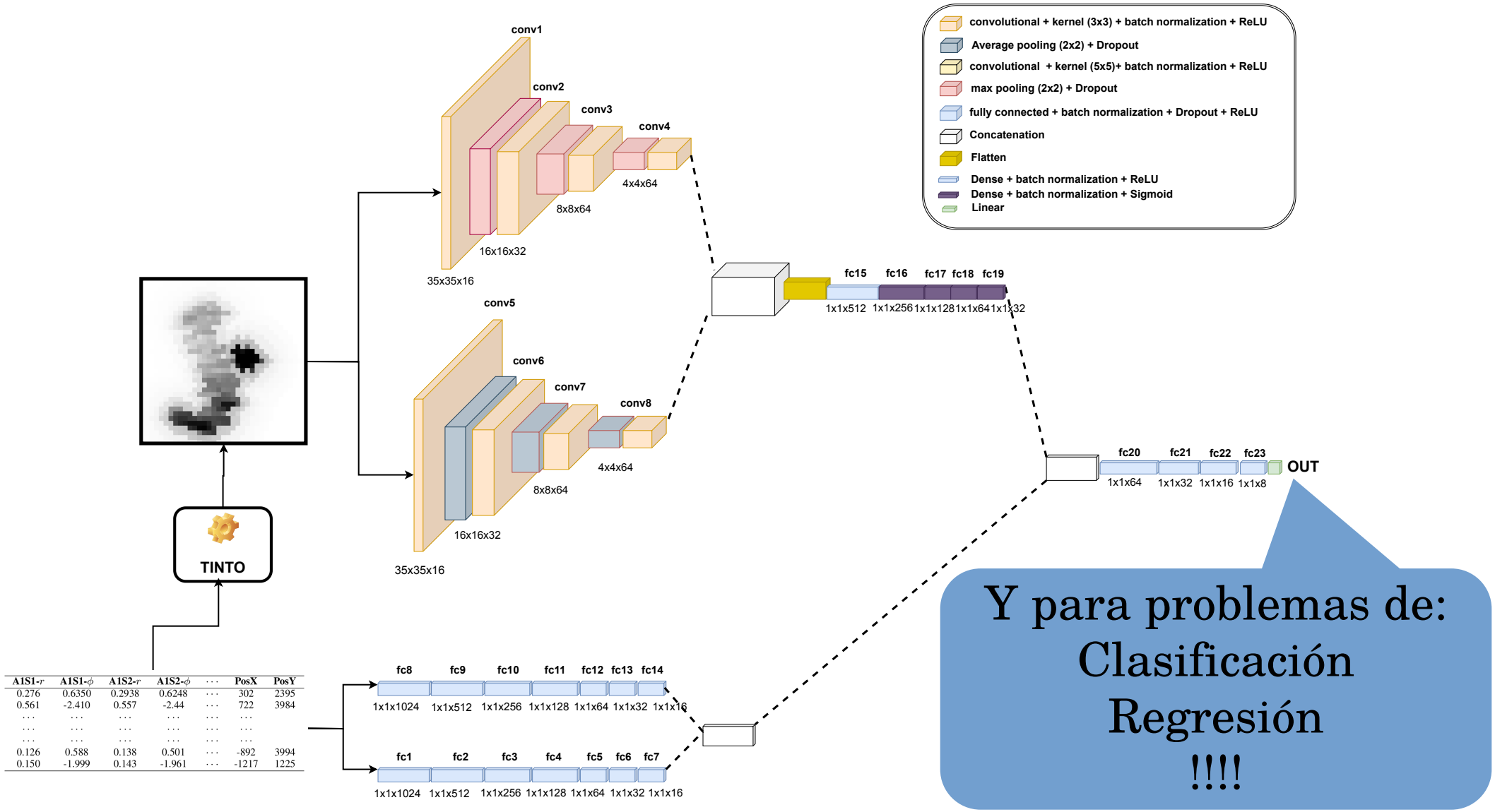
Pero no solamente CNNs

Hybrid Neural Network



A1S1-r	A1S1-ϕ	A1S2-r	A1S2-ϕ	...	PosX	PosY
0.276	0.6350	0.2938	0.6248	...	302	2395
0.561	-2.410	0.557	-2.44	...	722	3984
...	...	...	...	...	...	
...	...	...	...	...	...	
...	...	...	...	...	...	
0.126	0.588	0.138	0.501	...	-892	3994
0.150	-1.999	0.143	-1.961	...	-1217	1225

Hybrid Neural Network



A1S1-r	A1S1-ϕ	A1S2-r	A1S2-ϕ	···	PosX	PosY
0.276	0.6350	0.2938	0.6248	···	302	2395
0.561	-2.410	0.557	-2.44	···	722	3984
···	···	···	···	···	···	
···	···	···	···	···	···	
···	···	···	···	···	···	
0.126	0.588	0.138	0.501	···	-892	3994
0.150	-1.999	0.143	-1.961	···	-1217	1225

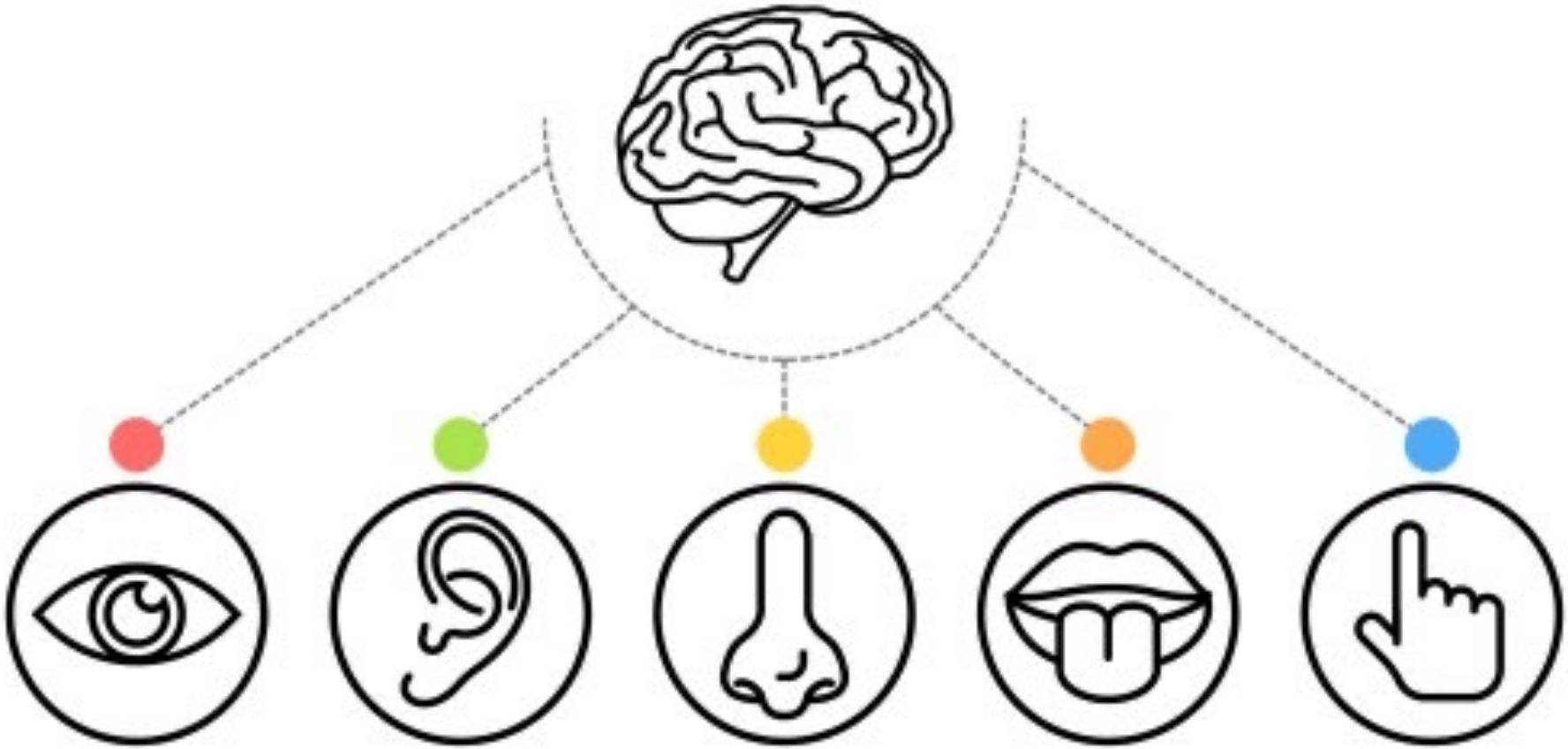
Y para problemas de:
Clasificación
Regresión
!!!!

Redes Multimodales

ETS de Ingeniería Informática

UNED

Basadas en el ser humano

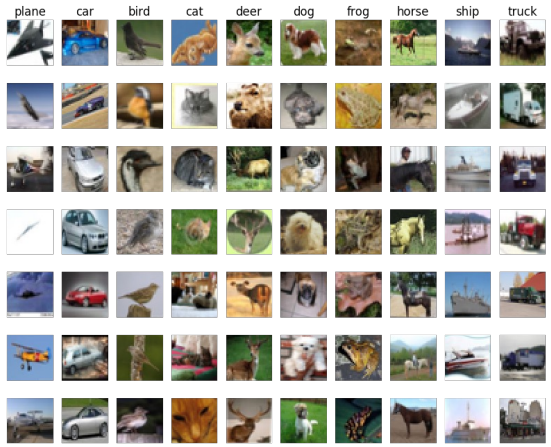


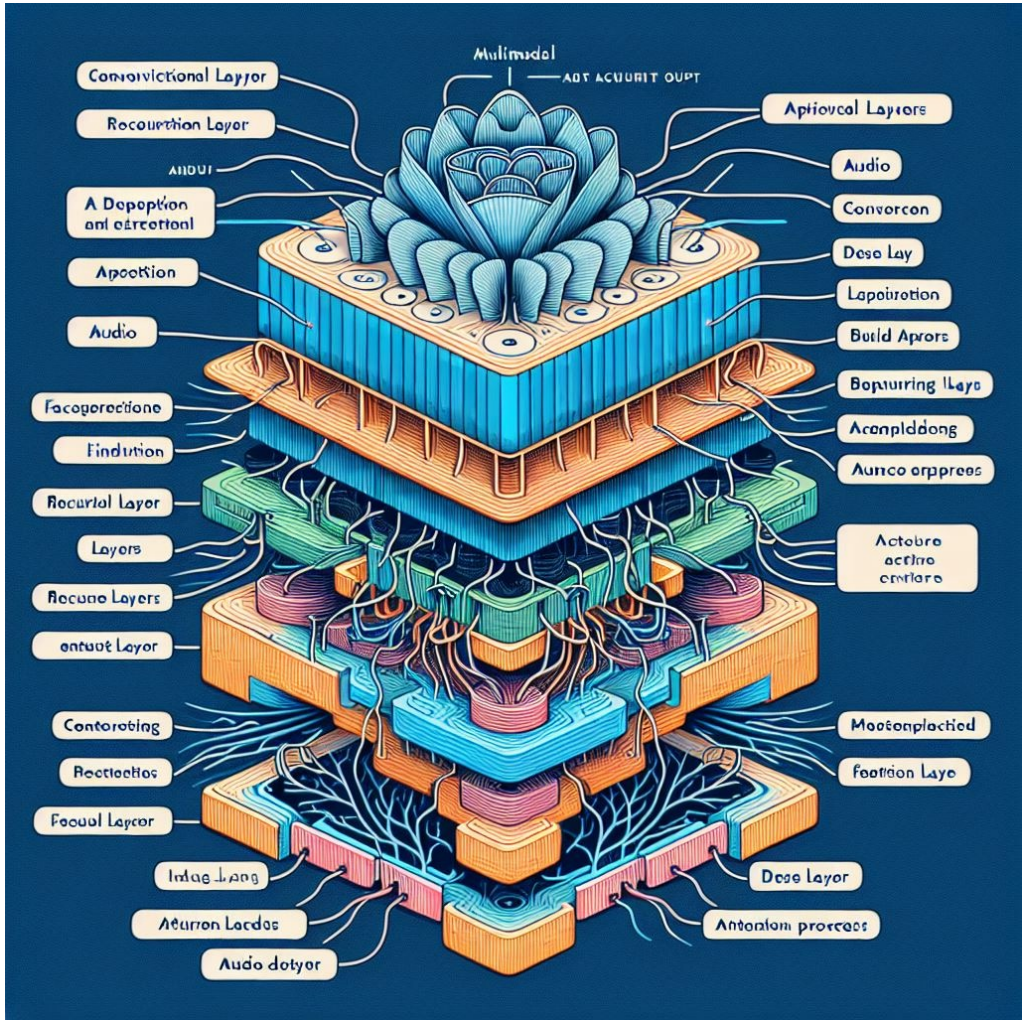
Basadas en el ser humano

Basadas en el ser humano

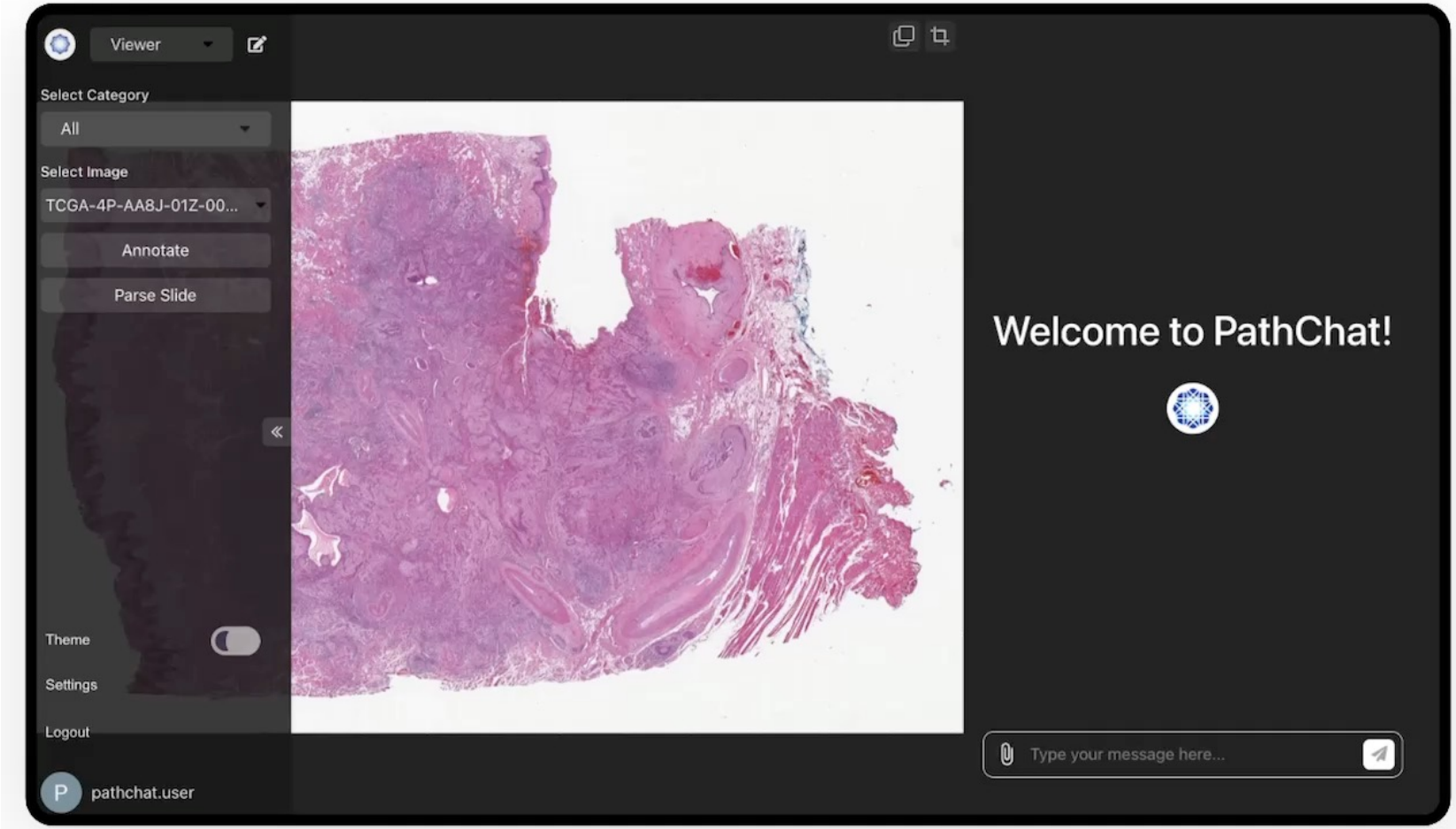
Basadas en el ser humano

Basadas en el ser humano

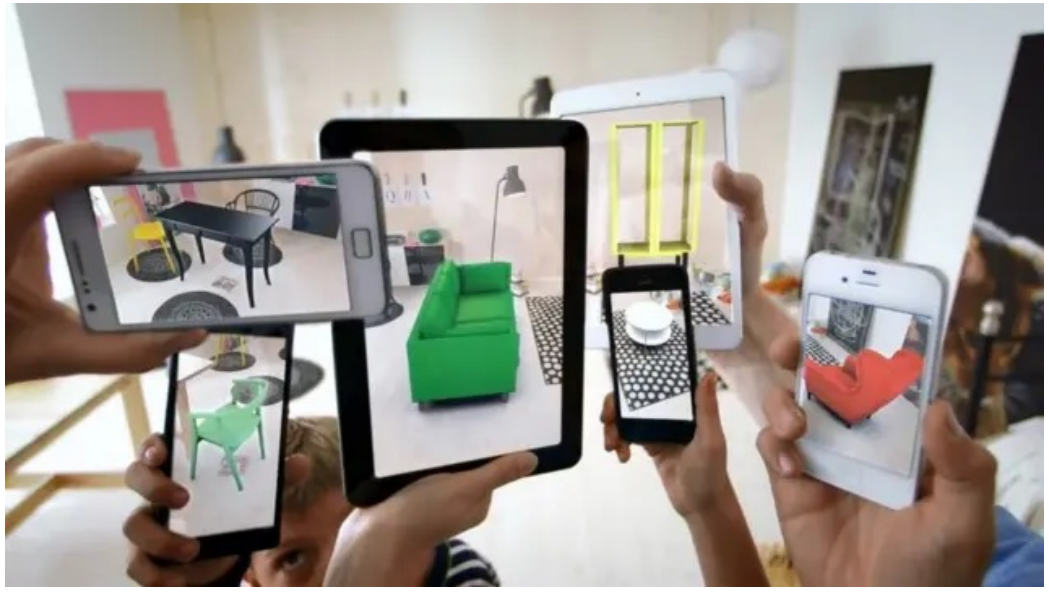


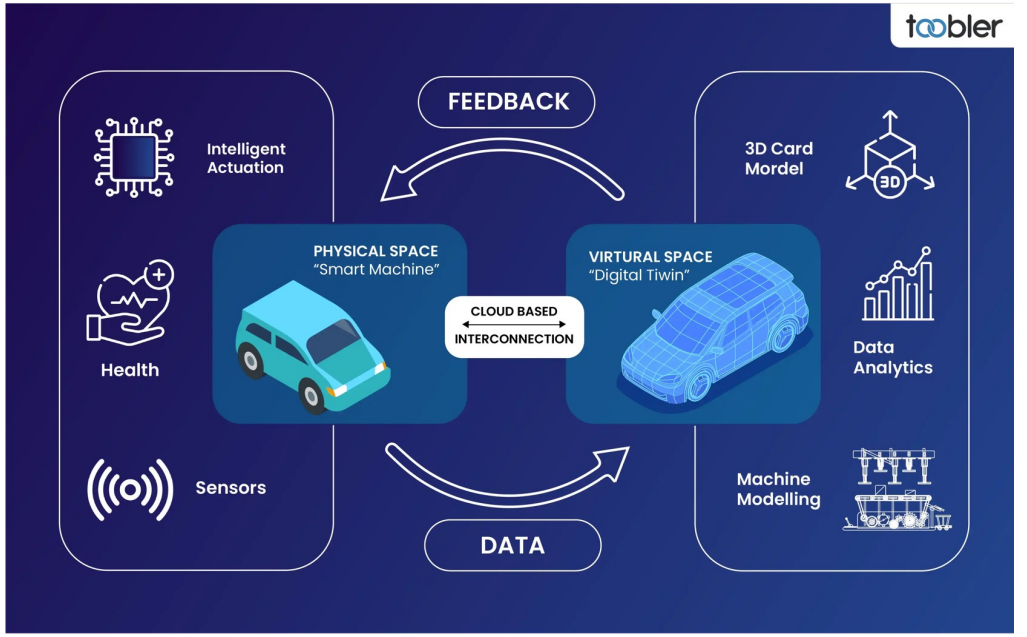


Por poner un ejemplo



Pero aplicado a todo





Pero sobre todo a...

Métodos de transformación

ETS de Ingeniería Informática

UNED

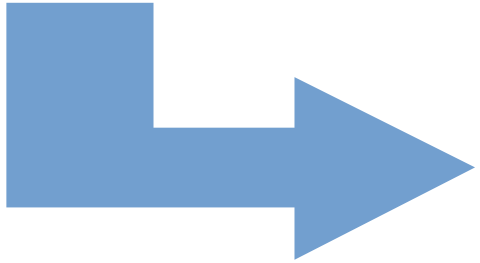
TINTOlib

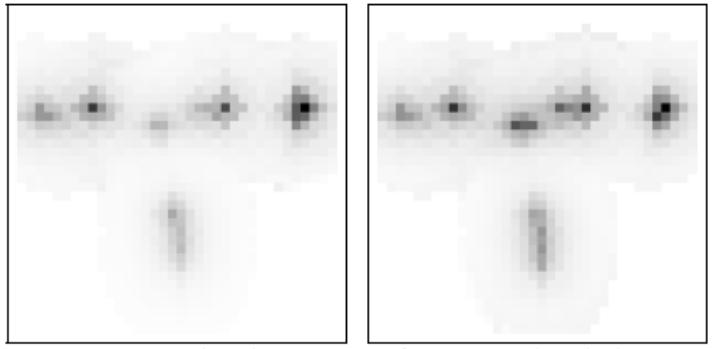
UNED

POLITÉCNICA

Transformación de datos

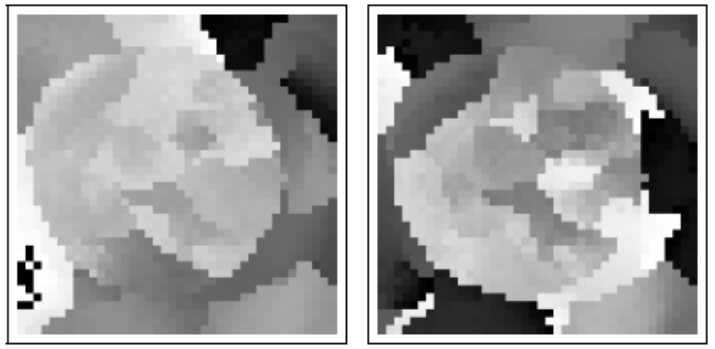
Be07	Be08	Be09	Be10	Be11	Sector
−65	−61	−74	−73	−67	1
−60	−57	−83	−62	−69	2
−66	−70	−78	−63	−73	3
...	...	...	...	...	...
−58	−66	−71	−73	−69	14
−60	−62	−73	−69	−57	15





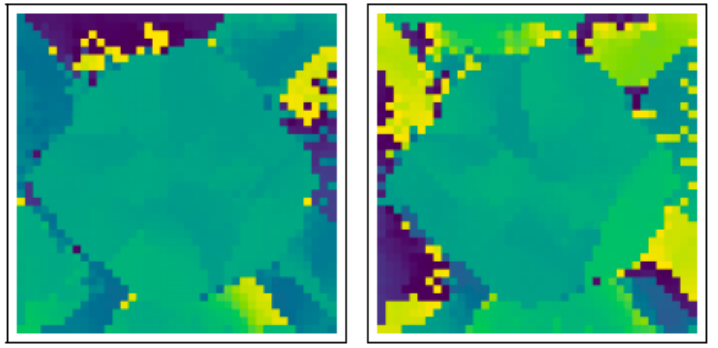
(a) TINTO - Sample 1.

(b) TINTO - Sample 50,000.



(c) IGTD - Sample 1.

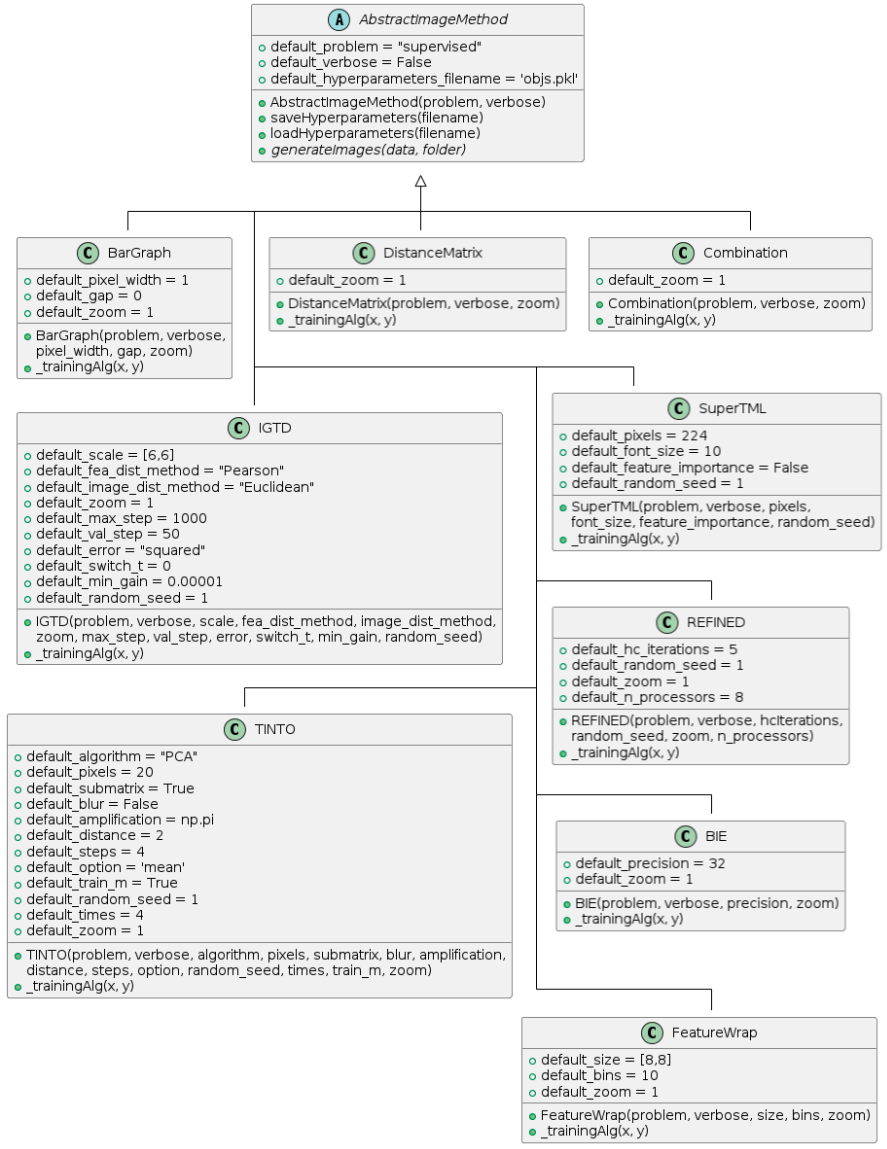
(d) IGTD - Sample 50,000.



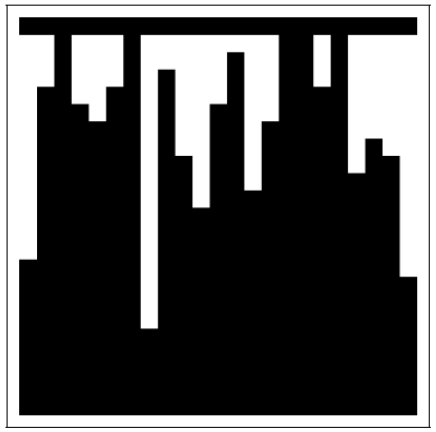
(e) REFINED - Sample 1.

(f) REFINED - Sample 50,000.

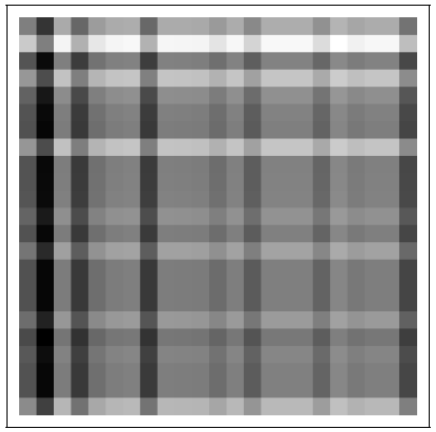
Métodos de transformación



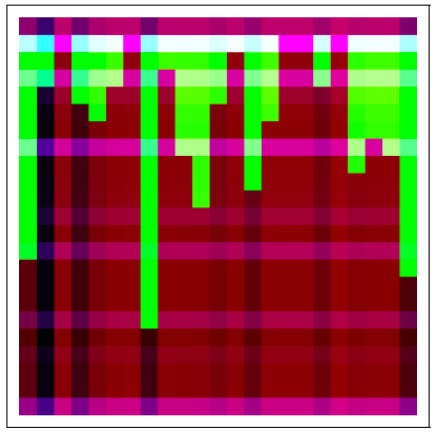
Métodos de transformación



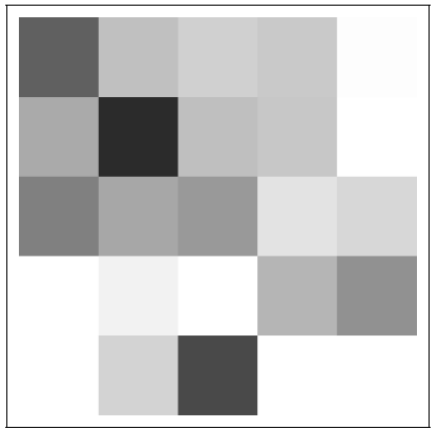
BarGraph



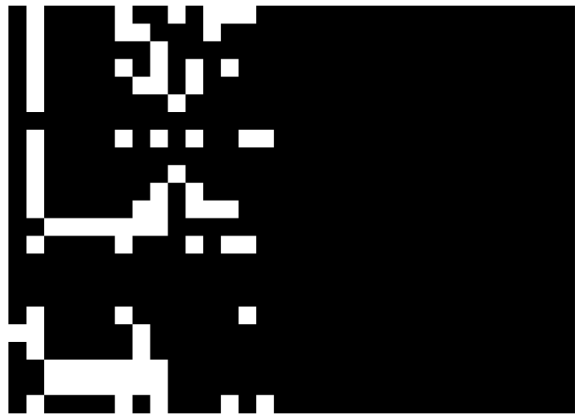
DistanceMatrix



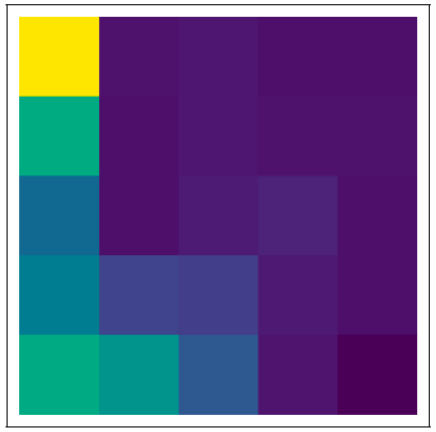
Combination



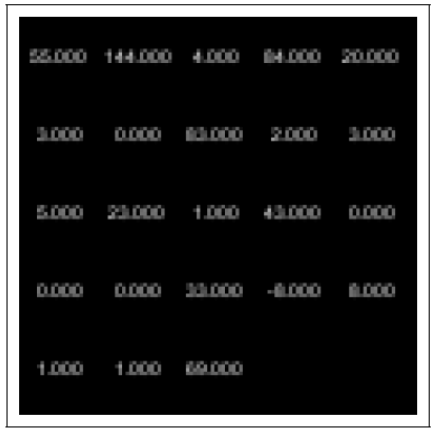
IGTD



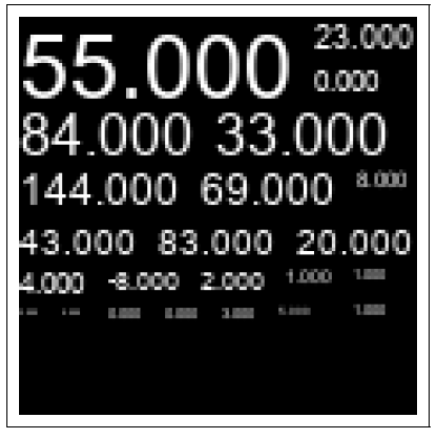
BIE



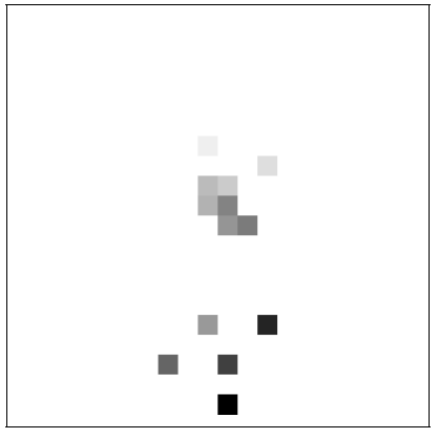
REFINED



SuperTML-EF



SuperTML-VF



TINTO

Feature Wrap

Problema de investigación abierto

Problema de investigación abierto

- Kadra *et al*. nombró datasets como "***last unconquered castle***" para modelos basados en Deep Neural Networks (DNN).
- La adaptación de las DNN a Datos Tabulares (TD) para tareas de inferencia o generación de datos sigue siendo un **gran desafío**.

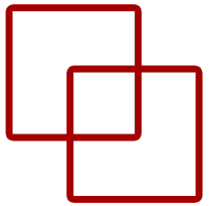
The references or papers are included as embedded links within the text.

Problema de investigación abierto

- Kadra *et al*. nombró datasets como "***last unconquered castle***" para modelos basados en Deep Neural Networks (DNN).
- La adaptación de las DNN a Datos Tabulares (TD) para tareas de inferencia o generación de datos sigue siendo un **gran desafío**.
- Vadim Borísov *et al*. hacen un benchmark de diferentes métodos/datasets entre ensembles vs. DNN y **ganan los ensembles**. De hecho, afirman:
  - "El progreso de la investigación sobre modelos competitivos de aprendizaje profundo para datos tabulares **se está estancando**".
  - "Es un **área de investigación abierta**".

The references or papers are included as embedded links within the text.

Problema de investigación abierto



- Kadra *et al*. nombró datasets como "***last unconquered castle***" para modelos basados en Deep Neural Networks (DNN).
- La adaptación de las DNN a Datos Tabulares (TD) para tareas de inferencia o generación de datos sigue siendo un **gran desafío**.
- Vadim Borísov *et al*. hacen un benchmark de diferentes métodos/datasets entre ensembles vs. DNN y **ganan los ensembles**. De hecho, afirman:
  - "El progreso de la investigación sobre modelos competitivos de aprendizaje profundo para datos tabulares **se está estancando**".
  - "Es un **área de investigación abierta**".
- Shwartz Ziv & Armon en su artículo *"Tabular data: Deep learning is not all you need"*, compararon enfoques de DNN vs. árbol de decisión de aumento de gradiente (GBDT). **Los GBDT cuestionaron los DNN**, lo que concluyeron que el modelado de datos tabulares utilizando DNN sigue siendo un **problema de investigación abierto**.

The references or papers are included as embedded links within the text.

Problema de investigación abierto

- Kadra *et al*. nombró datasets como "***last unconquered castle***" para modelos basados en Deep Neural Networks (DNN).
- La adaptación de las DNN a Datos Tabulares (TD) para tareas de inferencia o generación de datos sigue siendo un **gran desafío**.
- Vadim Borísov *et al*. hacen un benchmark de diferentes métodos/datasets entre ensembles vs. DNN y **ganan los ensembles**. De hecho, afirman:
  - "El progreso de la investigación sobre modelos competitivos de aprendizaje profundo para datos tabulares **se está estancando**".
  - "Es un **área de investigación abierta**".
- Shwartz Ziv & Armon en su artículo *"Tabular data: Deep learning is not all you need"*, compararon enfoques de DNN vs. árbol de decisión de aumento de gradiente (GBDT). **Los GBDT cuestionaron los DNN**, lo que concluyeron que el modelado de datos tabulares utilizando DNN sigue siendo un **problema de investigación abierto**.
- Las implementaciones exitosas de aplicaciones basadas en datos requieren resolver varias tareas, entre las cuales identificamos **3 desafíos centrales**: **(1) Inferencia; (2) Generación de datos; y (3) Interpretabilidad.**
  - La **tarea más crucial es la inferencia**, que se ocupa de hacer predicciones basadas en observaciones pasadas.

The references or papers are included as embedded links within the text.

Métodos no paramétricos

ETS de Ingeniería Informática

UNED

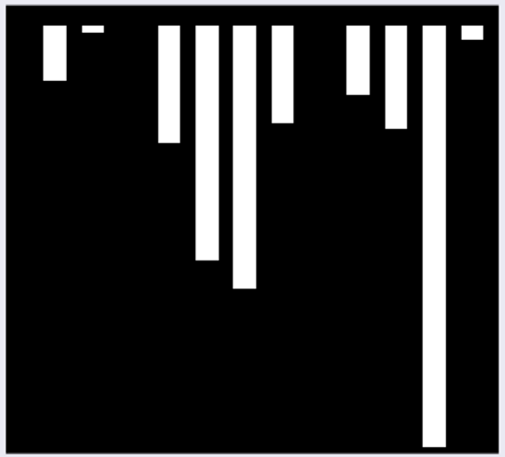
BarGraph

BarGraph

INPUT

0.00632	18.00	2.31	0	0.538	6.575	65.20	4.09	1	296.0	15.3	396	4.98

OUPUT

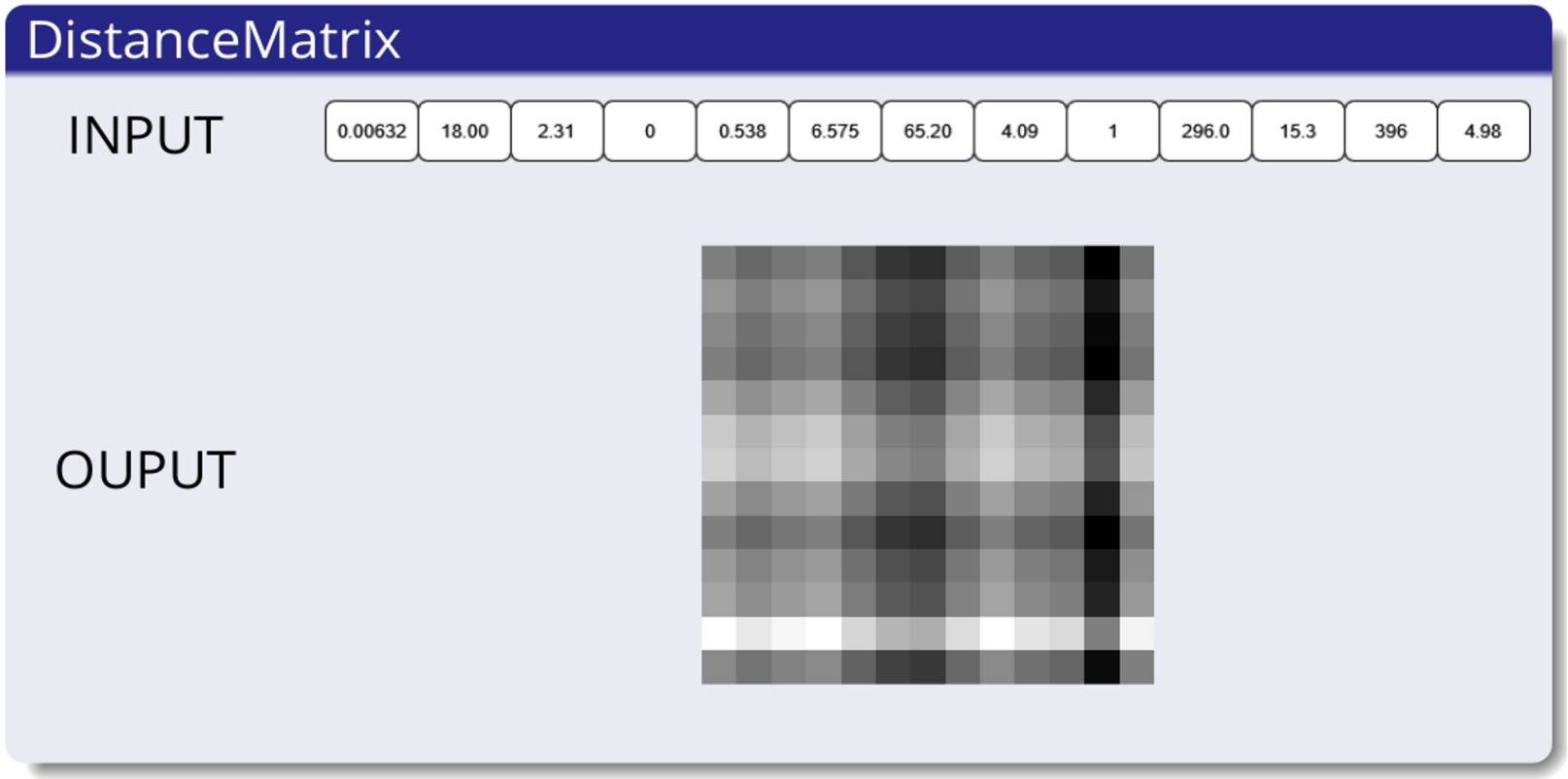


BarGraph

Parameters	Description	Default value	Valid values
`problem`	Defines the problem type for grouping images.	'supervised'	['supervised', 'unsupervised', 'regression']
`pixel_width`	The width (in pixels) for each column.	1	integer
`gap`	The separation (in pixels) between each column.	0	integer
`zoom`	Factor to scale the saved image size	1	int
`verbose`	Show in terminal the execution.	False	[True, False]

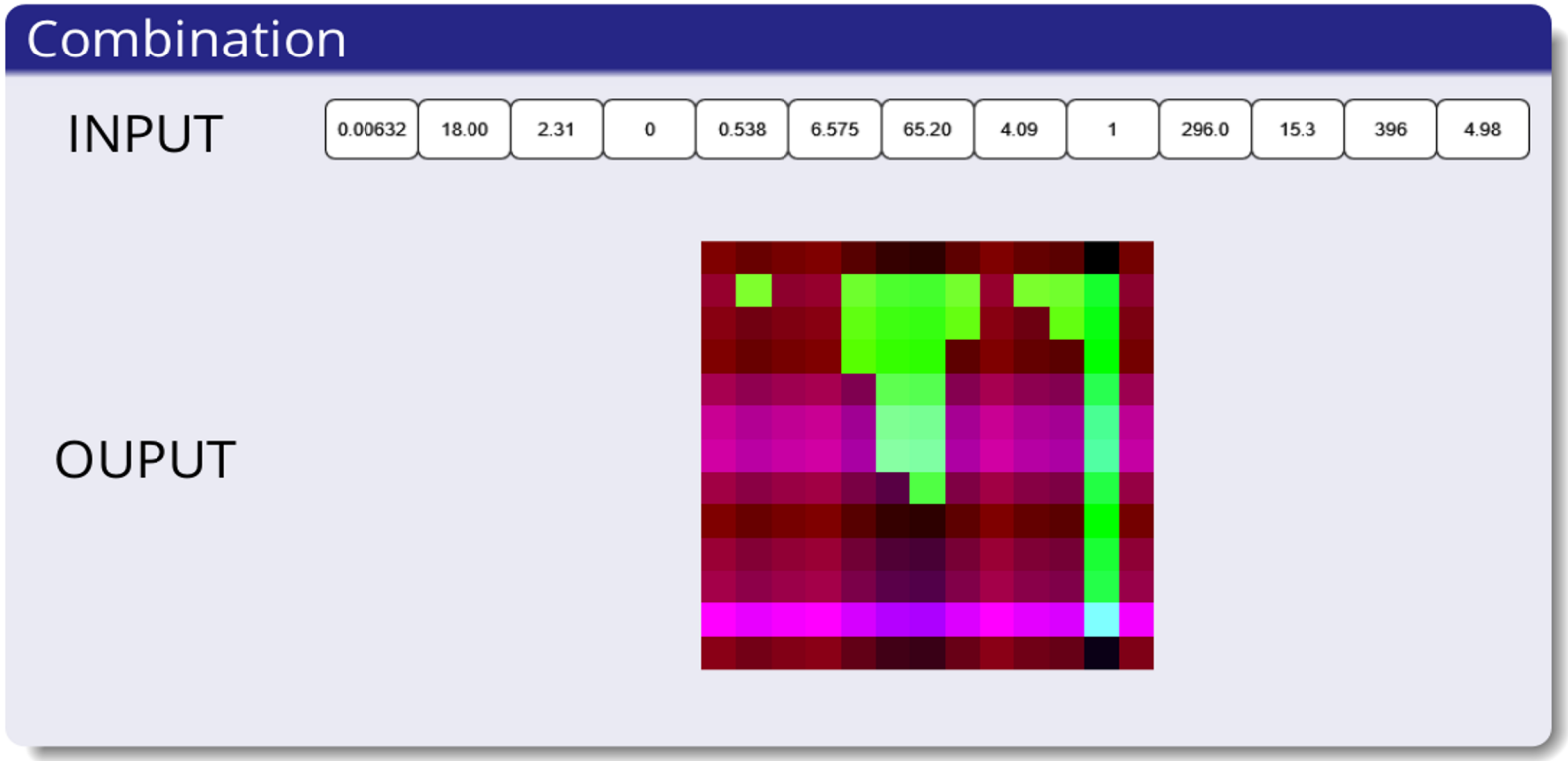
DistanceMatrix

Parameter	Description	Default	Valid Values
`zoom`	Scale factor for saved image size	1	int
`random_seed`	Seed for reproducibility	1	int
`verbose`	Show execution details	False	[True, False]



Combination

Parameter	Description	Default	Valid Values
zoom	Scale factor for saved image size	1	int
random_seed	Seed for reproducibility	1	int
verbose	Show execution details	False	[True, False]



SuperTML

Parameters	Description	Default value	Valid values
`problem`	Defines how images are grouped	'supervised'	['supervised', 'unsupervised', 'regression']
`image_pixels`	Number of pixels per side (total pixels = pixels × pixels)	224	integer
`feature_importance`	If False, uses equal font sizes (SuperTML-EF); if True, font size is proportional to feature importance (SuperTML-VF)	False	[True, False]
`font_size`	Font size used to render text on images	10	integer
`random_seed`	Seed for reproducibility	1	integer
`verbose`	Show execution details	False	[True, False]

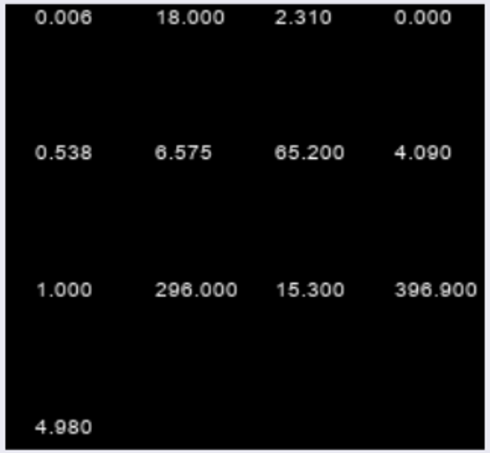
SuperTML

SuperTML

INPUT

0.00632	18.00	2.31	0	0.538	6.575	65.20	4.09	1	296.0	15.3	396	4.98

OUPUT



FeatureWrap

Parameters	Description	Default value	Valid values
problem	Defines how images are grouped	'supervised'	['supervised', 'unsupervised', 'regression']
size	Image dimensions in pixels (rows × columns)	[8,8]	[int, int]
bins	Number of bins for grouping numeric data	10	int
zoom	Factor to scale the saved image size	1	int
verbose	Show execution details	False	[True, False]

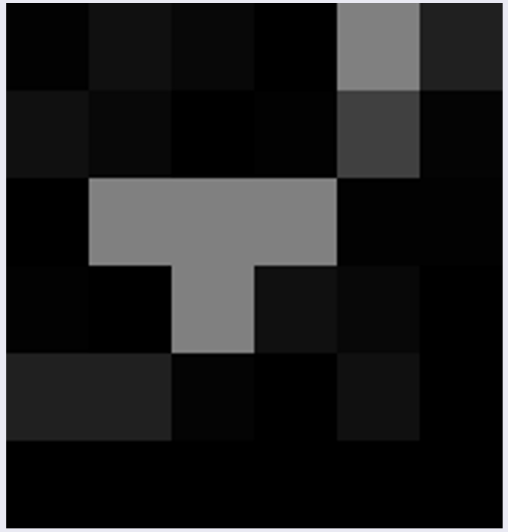
FeatureWrap

FeatureWrap

INPUT

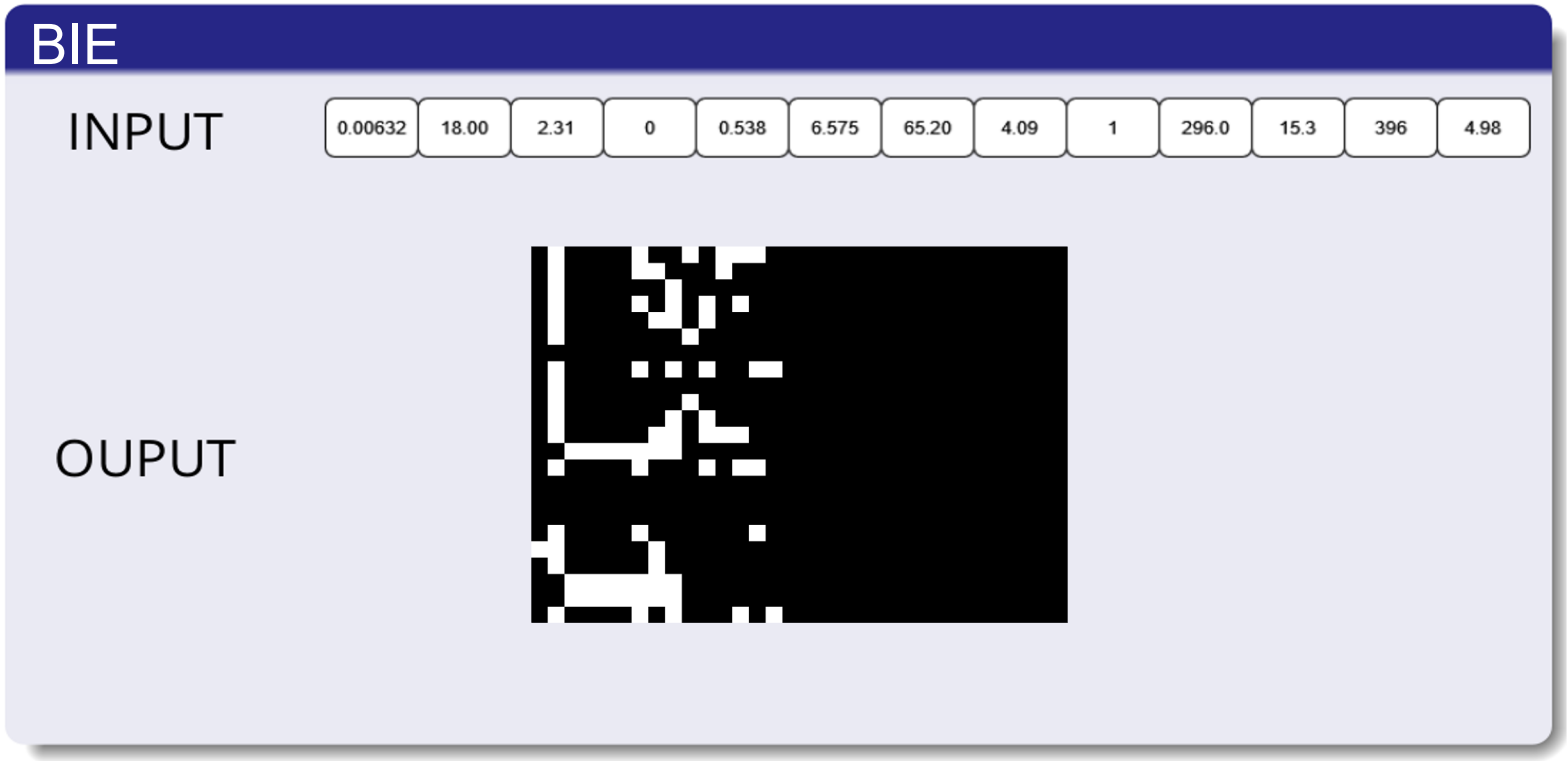
0.00632	18.00	2.31	0	0.538	6.575	65.20	4.09	1	296.0	15.3	396	4.98

OUPUT



BIE

Parameters	Description	Default value	Valid values
`problem`	The type of problem, this will define how the images are grouped.	'supervised'	['supervised', 'unsupervised', 'regression']
`precision`	Number of bits used to represent each feature.	32	[32, 64]
`zoom`	Factor to scale the saved image size	1	Positive integer
`verbose`	Show in terminal the execution.	False	[True, False]



Métodos paramétricos

ETS de Ingeniería Informática

UNED

TINTO

Parameter	Description	Default	Valid Values
problem	Defines the problem type for grouping images	'supervised'	['supervised', 'unsupervised', 'regression']
scale	Image size (rows × columns); must be ≥ number of features	[6, 6]	[int, int]
fea_dist_method	Method to evaluate feature similarity	'Pearson'	['Pearson', 'Spearman', 'set', 'Euclidean']
image_dist_method	Method to calculate distances	'Euclidean'	['Euclidean', 'Manhattan']
max_step	Max steps before algorithm stops if not converged	1000	int
val_step	Steps between convergence checks	50	int
error	Function to evaluate the difference between feature distance ranking and pixel distance ranking	'squared'	['squared', 'abs']
switch_t	Threshold to decide when to switch elements	0	int
min_gain	Minimum improvement needed to continue; else, algorithm stops	0.00001	float
zoom	Scale factor for saved image size	1	int
random_seed	Seed for reproducibility	1	int
verbose	Show execution details	False	[True, False]

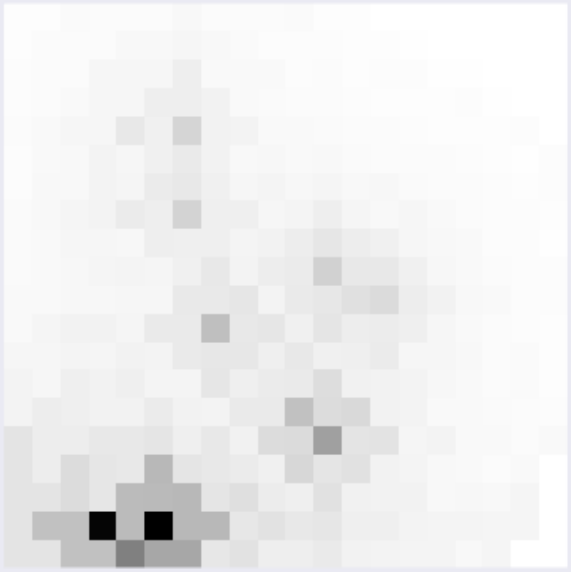
TINTO

TINTO

INPUT

0.00632	18.00	2.31	0	0.538	6.575	65.20	4.09	1	296.0	15.3	396	4.98

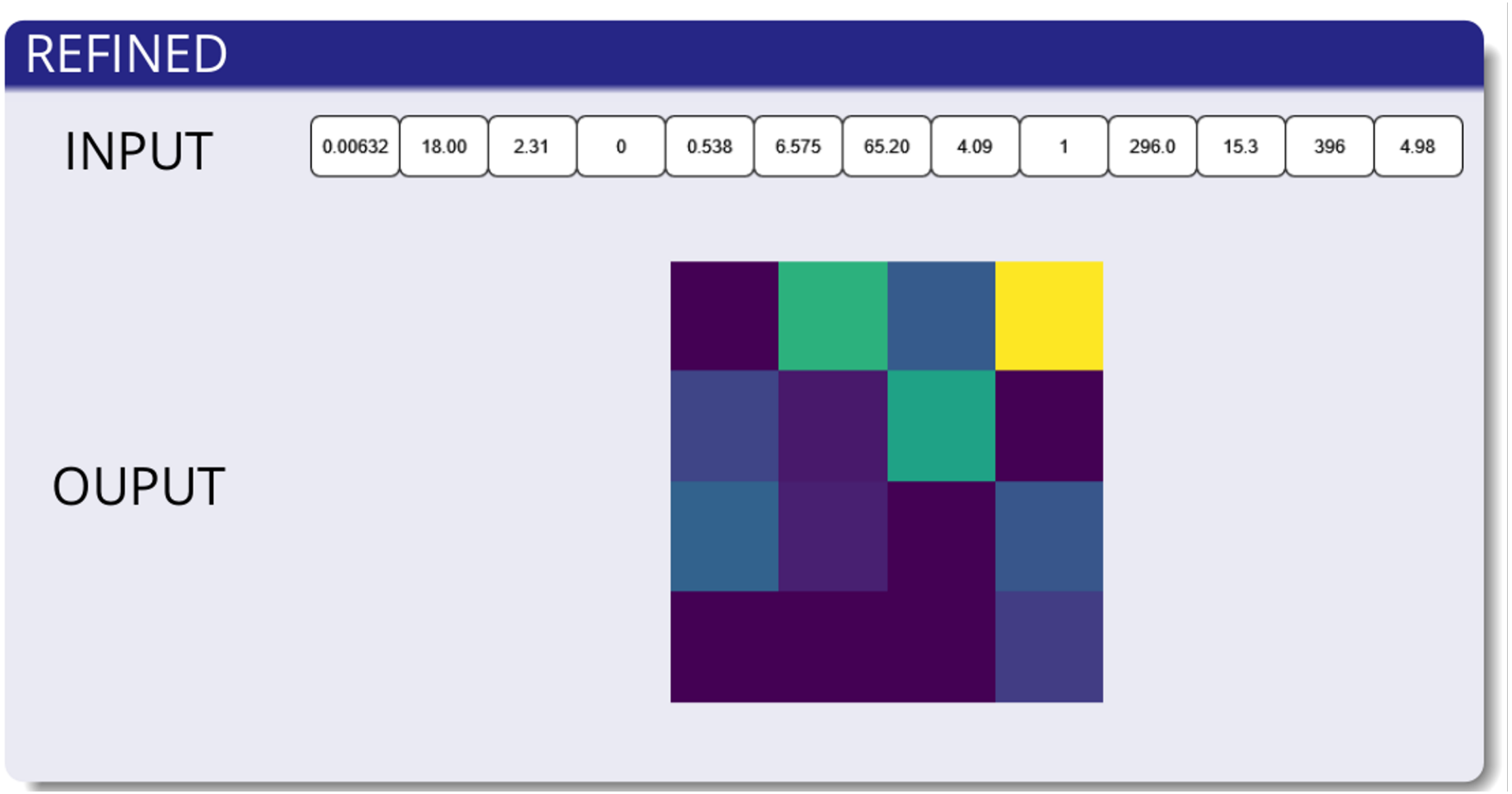
OUPUT



REFINED

Parameter	Description	Default	Valid Values
`problem`	Defines the problem type for grouping images	'supervised'	['supervised', 'unsupervised', 'regression']
`n_processors`	Number of processors to use (must be ≥ 2)	8	int (≥ 2)
`hcIterations`	Iterations of the hill climbing algorithm	5	int
`zoom`	Factor to scale the saved image size	1	int
`random_seed`	Seed for reproducibility	1	int
`verbose`	Show execution details in terminal	False	[True, False]

REFINED


REFINED
INPUT
0.00632
18.00
2.31
0
0.538
6.575
65.20
4.09
1
296.0
15.3
396
4.98
OUPUT

IGTD

Parameters	Description	Default value	Valid values
`problem`	Defines the problem type for grouping images	'supervised'	['supervised', 'unsupervised', 'regression']
`algorithm`	Chooses the dimensionality reduction algorithm	PCA	[PCA, t-SNE]
`pixels`	Sets the image size by specifying pixels per side (total pixels = pixels × pixels)	20	integer
`submatrix`	Determines if a submatrix is used for blurring	True	[True, False]
`blur`	Enables or disables blurring	False	[True, False]
`amplification`	Only with blur=true, blurring amplification	`np.pi`	float
`distance`	Only with blur=true, blurring distance in pixels	2	integer
`steps`	Only with blur=true, blurring steps	4	integer
`option`	Only with blur=true, method for handling overlapping pixels	mean	[mean, maximum]
`times`	Only with algorithm=t-SNE, times replication in t-SNE	4	integer
`zoom`	Factor to scale the saved image size	1	int
`random_seed`	Seed for reproducibility	1	integer
`verbose`	Show in terminal the execution	False	[True, False]

IGTD

IGTD

INPUT

0.00632	18.00	2.31	0	0.538	6.575	65.20	4.09	1	296.0	15.3	396	4.98

OUPUT

Arquitecturas con imágenes sintéticas

ETS de Ingeniería Informática

UNED

CNN

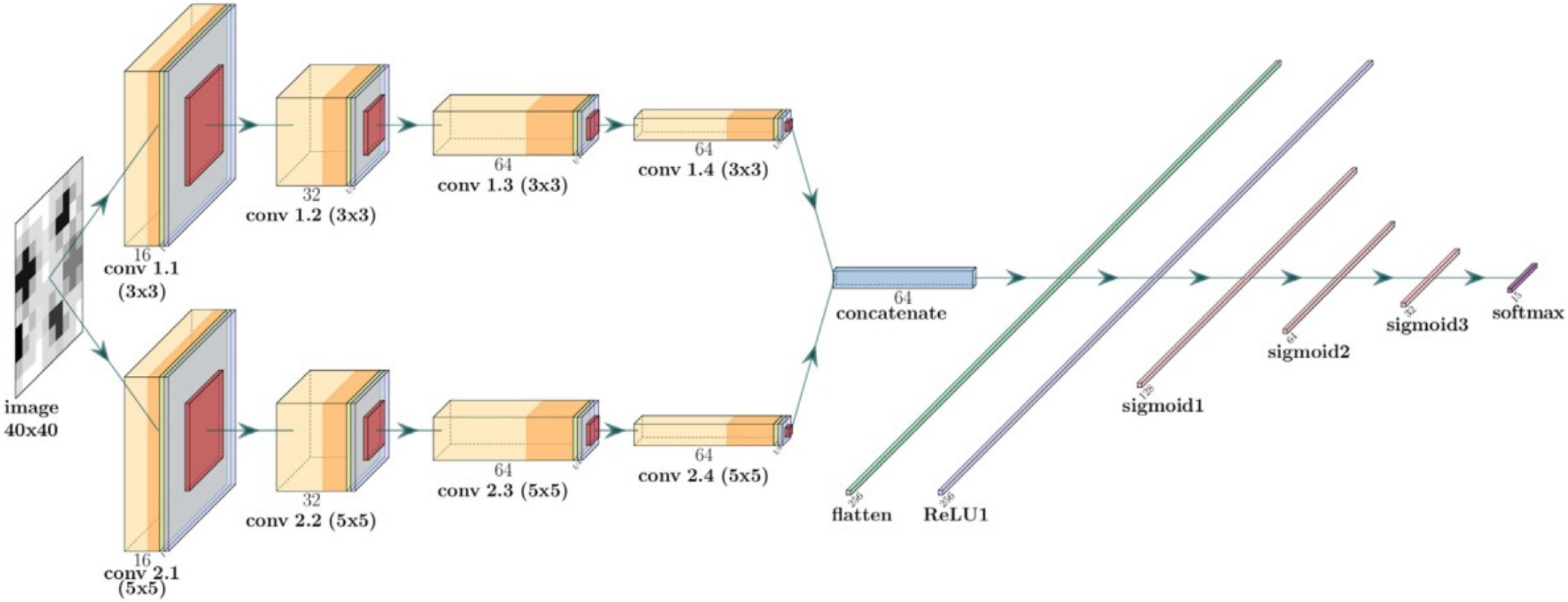
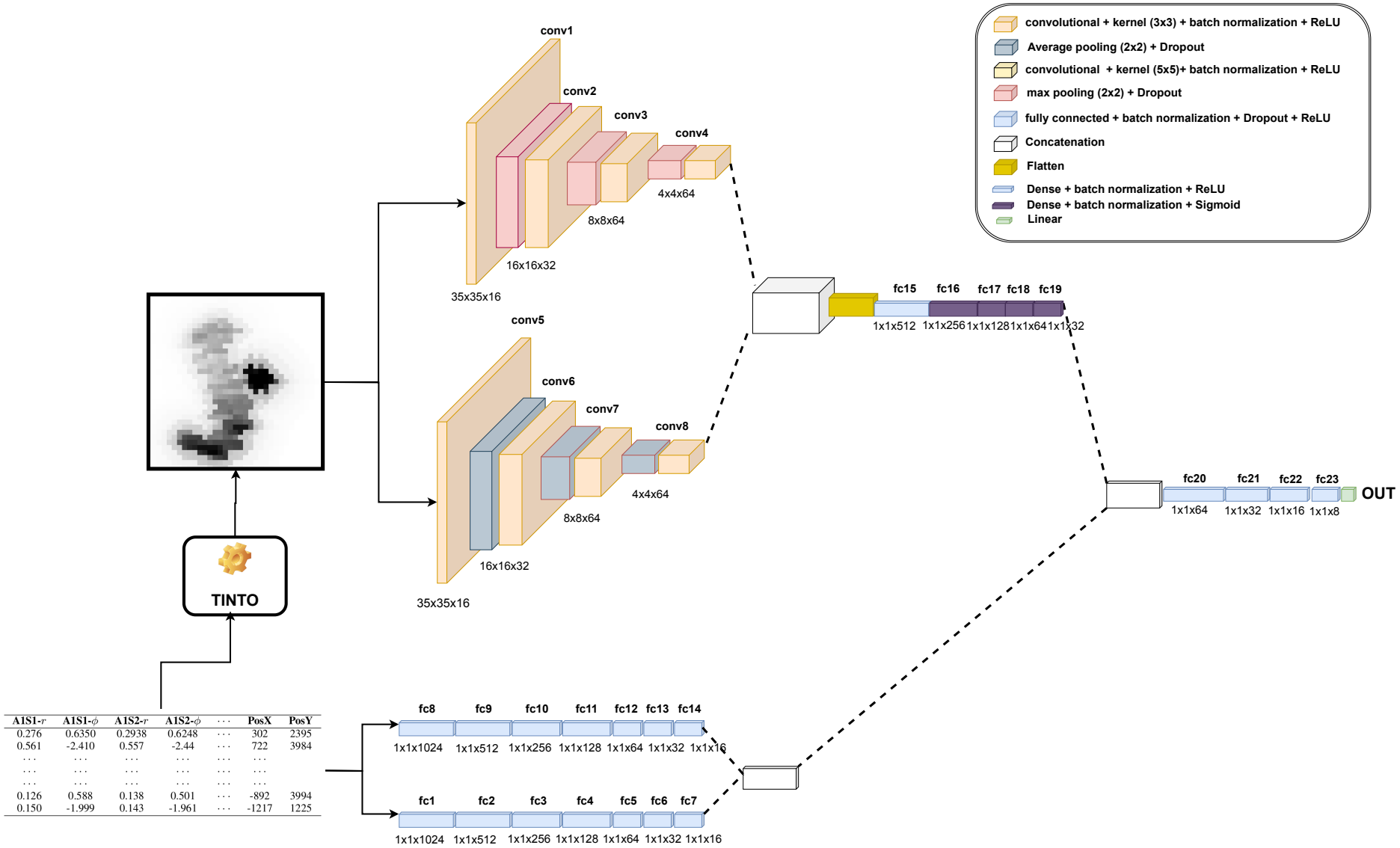


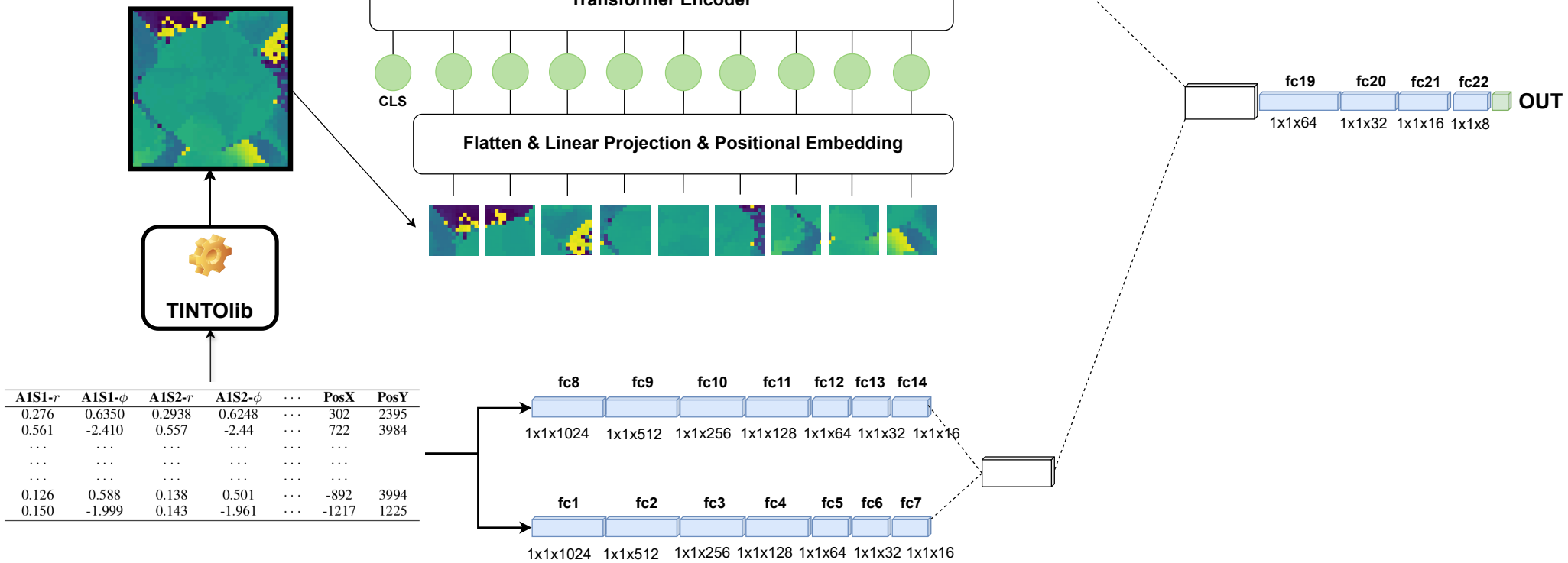
image
40x40
16
conv 1.1
(3x3)
32
conv 1.2 (3x3)
64
conv 1.3 (3x3)
64
conv 1.4 (3x3)
16
conv 2.1
(5x5)
32
conv 2.2 (5x5)
64
conv 2.3 (5x5)
64
conv 2.4 (5x5)
64
concatenate
flatten
ReLU1
sigmoid1
sigmoid2
sigmoid3
softmax

HyNN



A1S1-r	A1S1-ϕ	A1S2-r	A1S2-ϕ	...	PosX	PosY
0.276	0.6350	0.2938	0.6248	...	302	2395
0.561	-2.410	0.557	-2.44	...	722	3984
...	...	...	...	...	...	
...	...	...	...	...	...	
...	...	...	...	...	...	
0.126	0.588	0.138	0.501	...	-892	3994
0.150	-1.999	0.143	-1.961	...	-1217	1225

HyViT



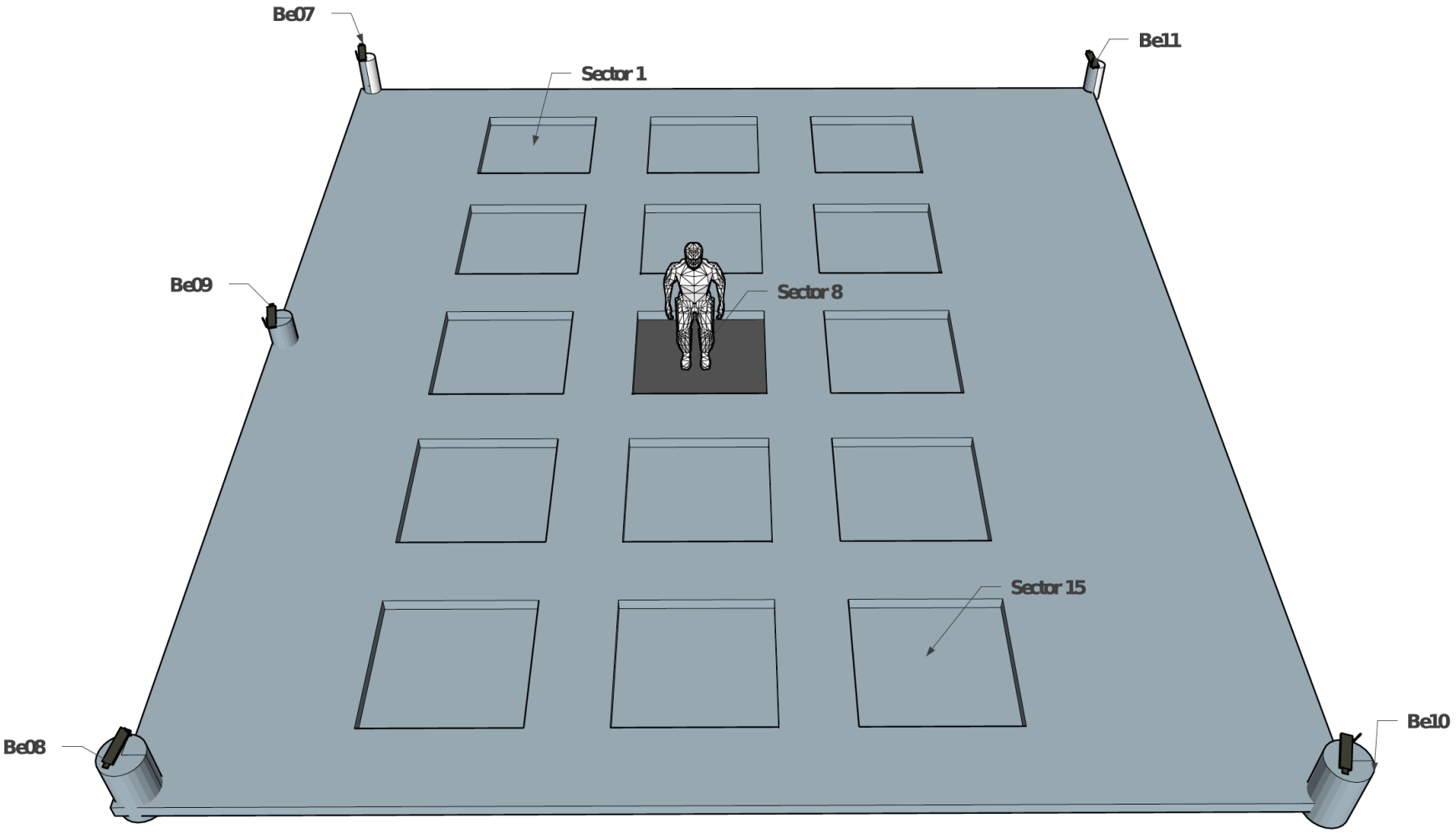
A1S1-r	A1S1-ϕ	A1S2-r	A1S2-ϕ	···	PosX	PosY
0.276	0.6350	0.2938	0.6248	···	302	2395
0.561	-2.410	0.557	-2.44	···	722	3984
···	···	···	···	···	···	
···	···	···	···	···	···	
···	···	···	···	···	···	
0.126	0.588	0.138	0.501	···	-892	3994
0.150	-1.999	0.143	-1.961	···	-1217	1225

Caso de uso 1: Localización en interiores con MIMO

ETS de Ingeniería Informática

UNED

Escenario



Talla-Chumpitaz, R., Castillo-Cara, M., Orozco-Barbosa, L., & García-Castro, R. (2023). A novel deep learning approach using blurring image techniques for Bluetooth-based indoor localisation. *Information Fusion*, *91*, 173-186.

Escenario

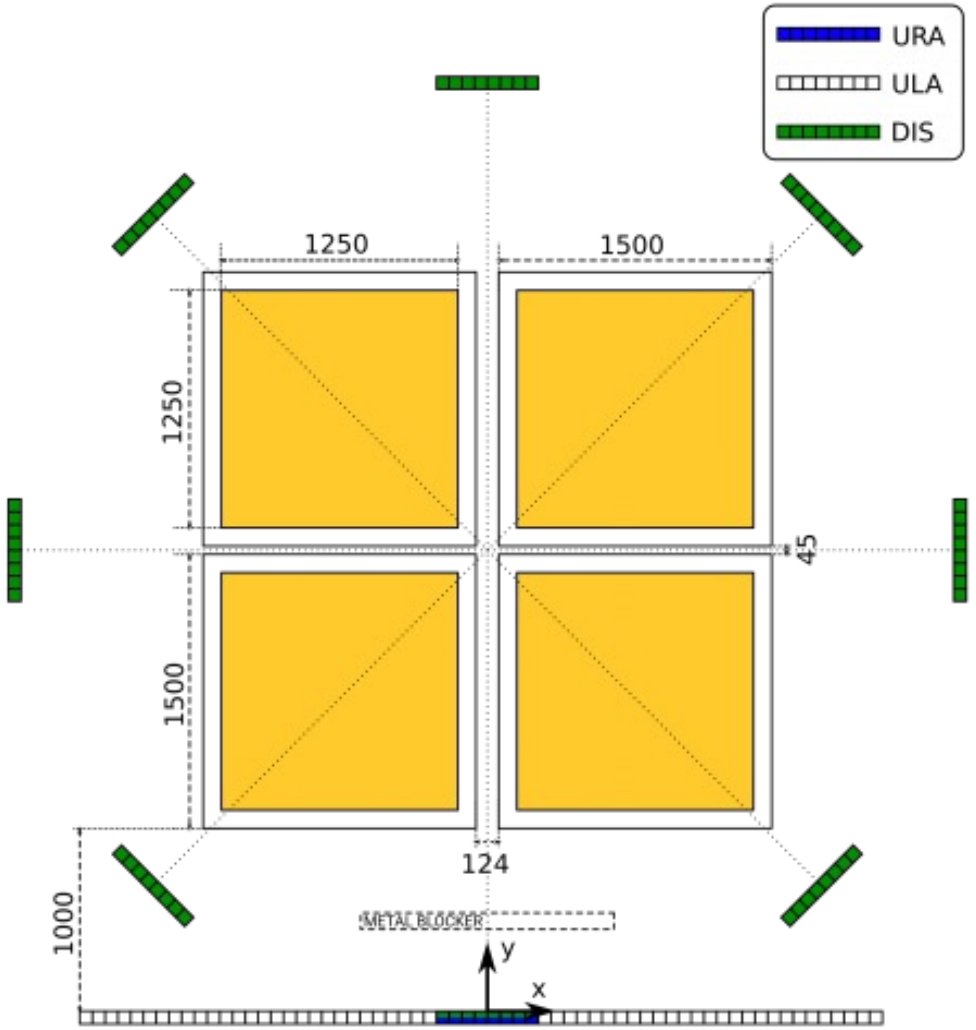
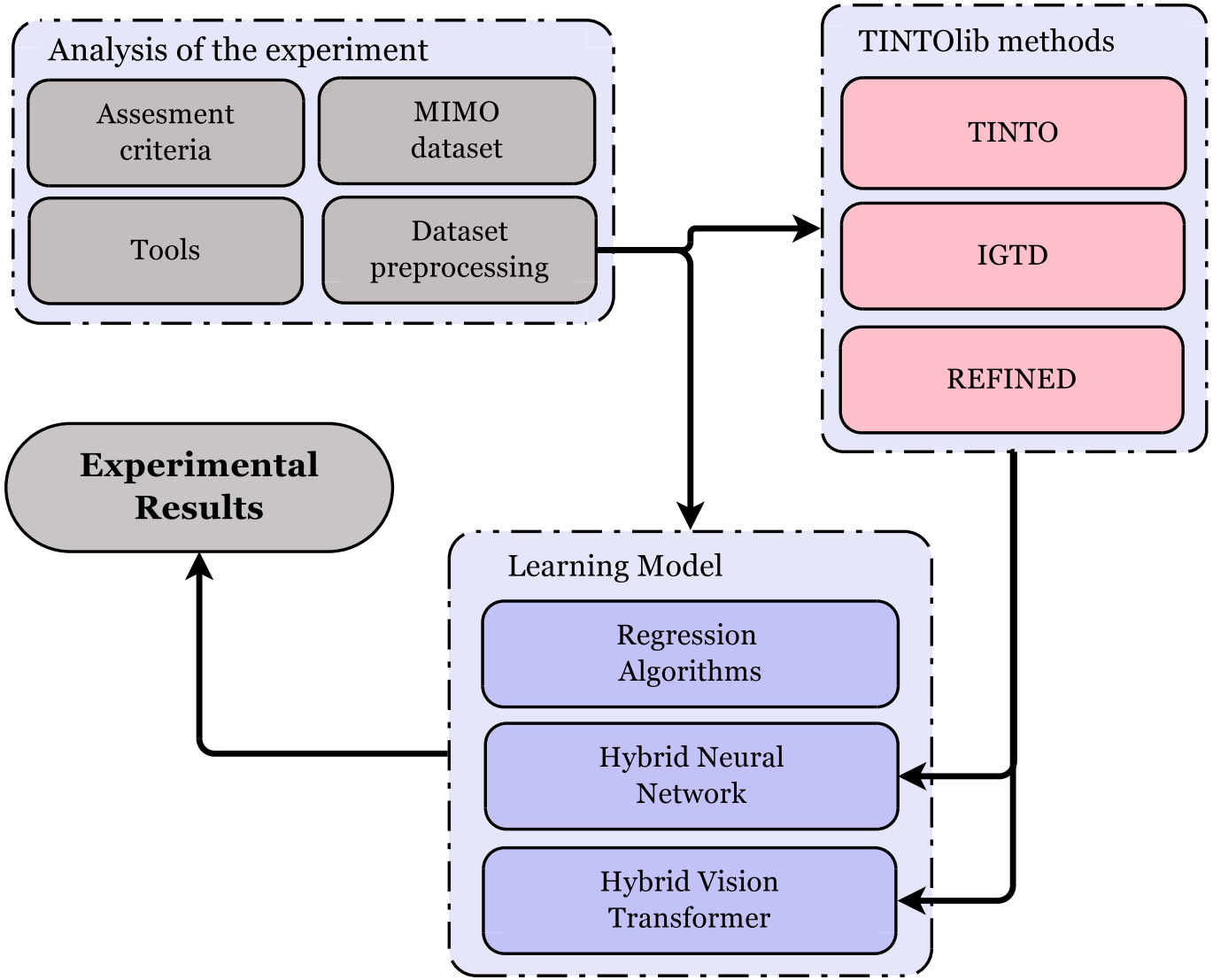


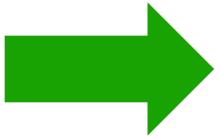
Figure taken from the paper: S. D. Bast, A. P. Guevara and S. Pollin, "CSI-based Positioning in Massive MIMO systems using Convolutional Neural Networks," 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium, 2020, pp. 1-5

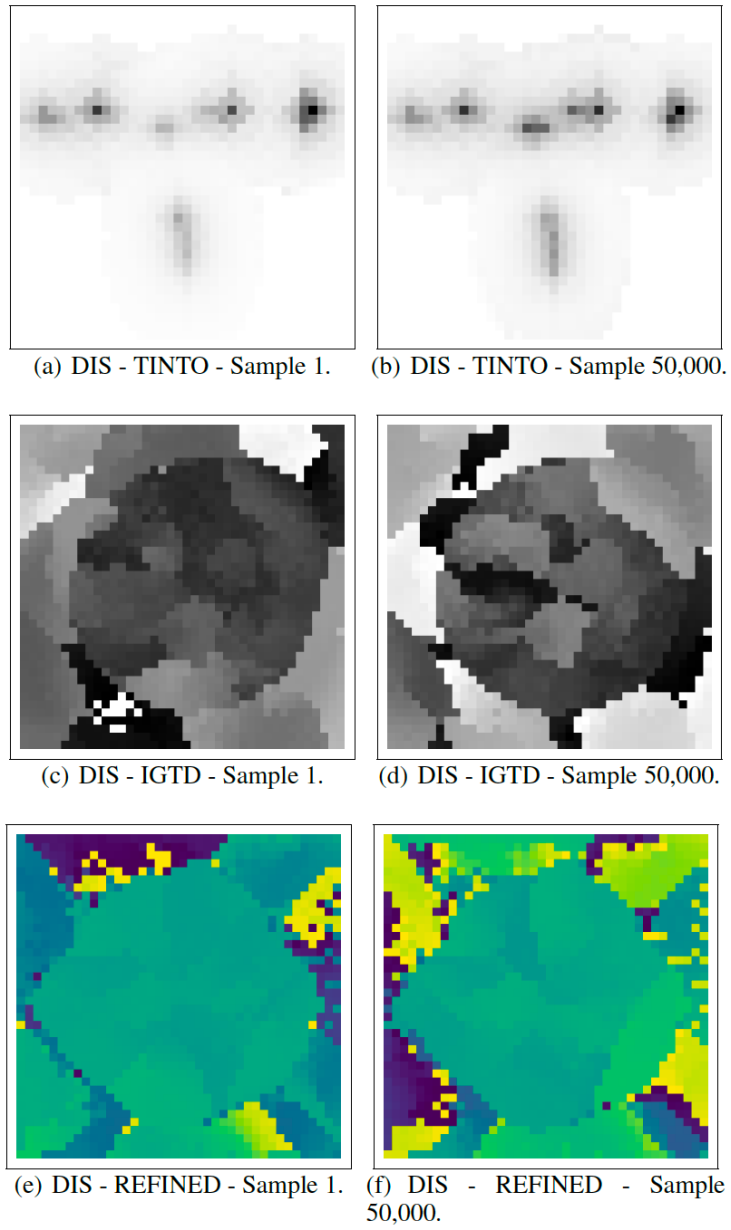
Metodología



Imágenes sintéticas

A1C1-m	A1C1-ϕ	A1C2-m	A1C2-ϕ	$\cdots$	PosX	PosY
0.276	0.6350	0.2938	0.6248	$\cdots$	302	2395
0.561	-2.410	0.557	-2.44	$\cdots$	722	3984
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	
0.126	0.588	0.138	0.501	$\cdots$	-892	3994
0.150	-1.999	0.143	-1.961	$\cdots$	-1217	1225

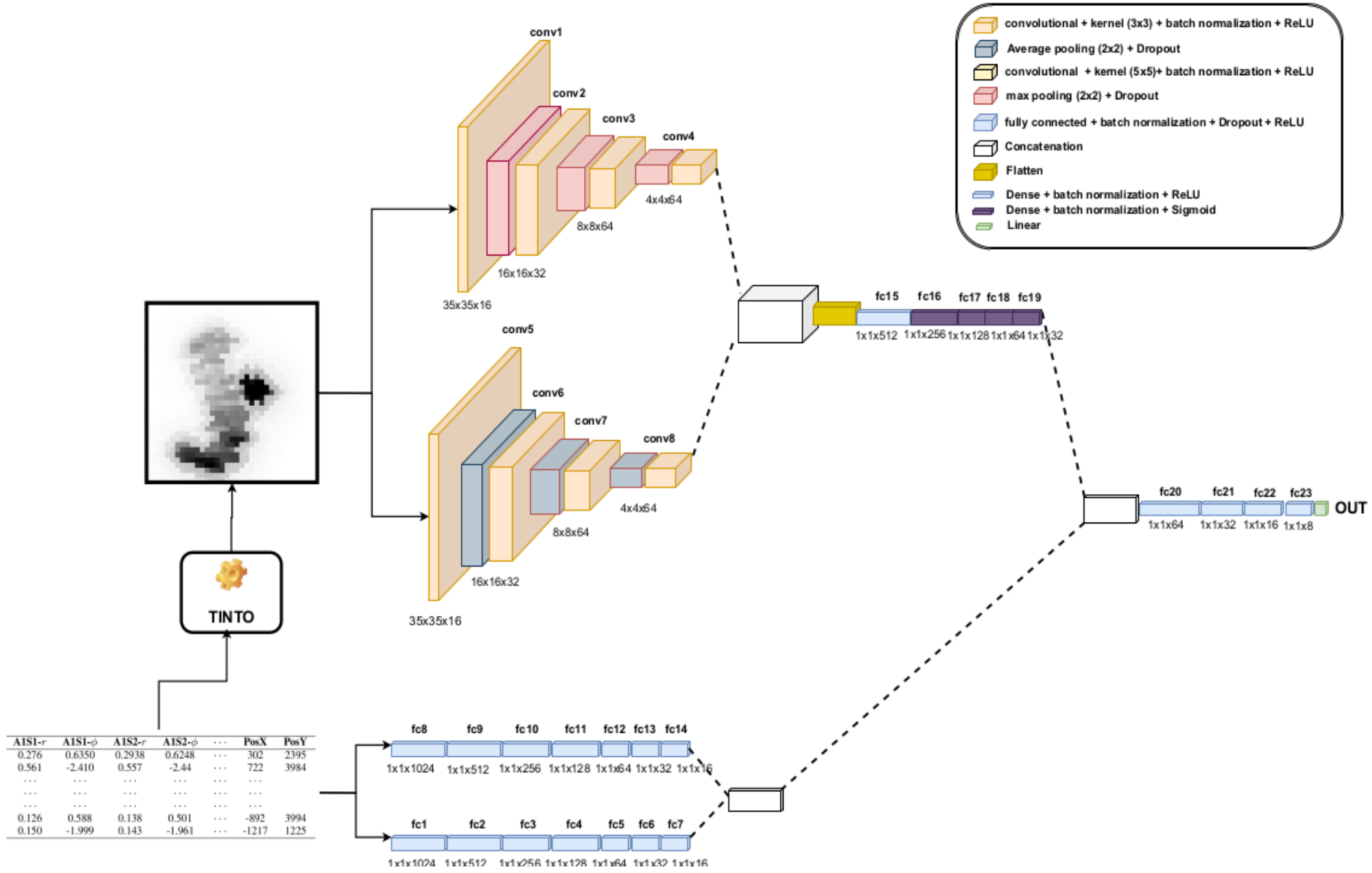




(a) DIS - TINTO - Sample 1. (b) DIS - TINTO - Sample 50,000. (c) DIS - IGTD - Sample 1. (d) DIS - IGTD - Sample 50,000. (e) DIS - REFINED - Sample 1. (f) DIS - REFINED - Sample 50,000.

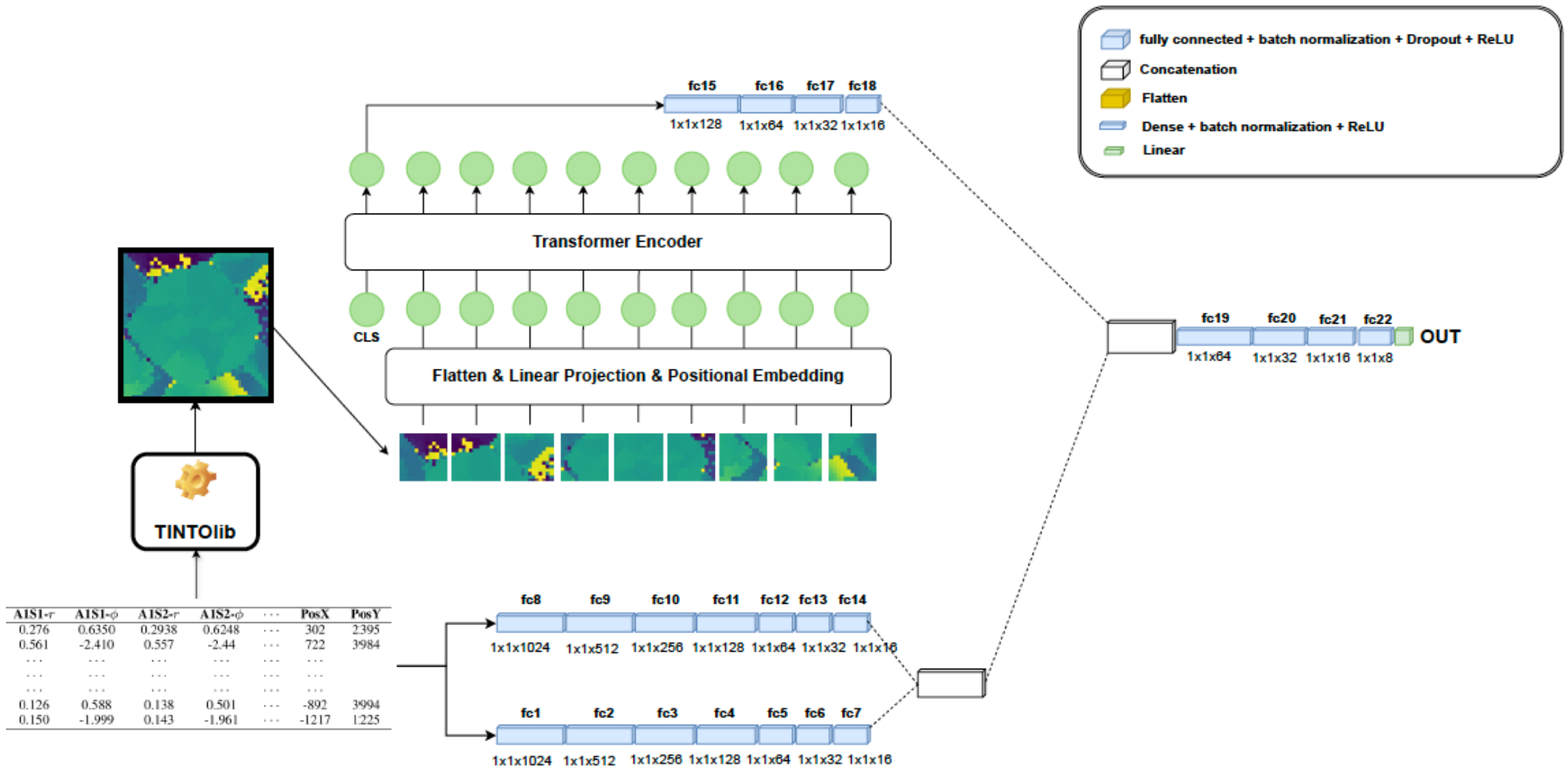
Figure 3. Synthetic image samples generated by TINTOlib for different samples in 8 antennas DIS scenario.

HyNN → CNN+MLP



A1S1-r	A1S1-ϕ	A1S2-r	A1S2-ϕ	···	PosX	PosY
0.276	0.6350	0.2938	0.6248	···	302	2395
0.561	-2.410	0.557	-2.44	···	722	3984
···	···	···	···	···	···	
···	···	···	···	···	···	
···	···	···	···	···	···	
0.126	0.588	0.138	0.501	···	-892	3994
0.150	-1.999	0.143	-1.961	···	-1217	1225

HyViT→ ViT+MLP



A1S1-r	A1S1-ϕ	A1S2-r	A1S2-ϕ	...	PosX	PosY
0.276	0.6350	0.2938	0.6248	...	302	2395
0.561	-2.410	0.557	-2.44	...	722	3984
...	...	...	...	...	...	
...	...	...	...	...	...	
...	...	...	...	...	...	
0.126	0.588	0.138	0.501	...	-892	3994
0.150	-1.999	0.143	-1.961	...	-1217	1225

Baseline Results

Table 2. RMSE (in mm) in validation (Val) and test split. Best results are shown in bold.

Algorithm	PosX		PosY	
	Val	Test	Val	Test
BR	226.05	225.00	251.43	255.54
ET	163.15	161.65	180.00	185.70
HGB	194.10	194.97	236.55	236.46
KNN	**110.50**	**110.54**	**133.70**	**140.16**
LiR	383.05	386.95	432.83	439.10
MLP	179.80	178.82	326.11	334.76
RF	226.09	225.18	251.37	255.62
RCV	383.04	386.94	432.80	439.06
XGB	178.41	180.03	202.45	201.66
LGB	194.14	194.15	231.19	232.89

Hybrid Neural Networks Results

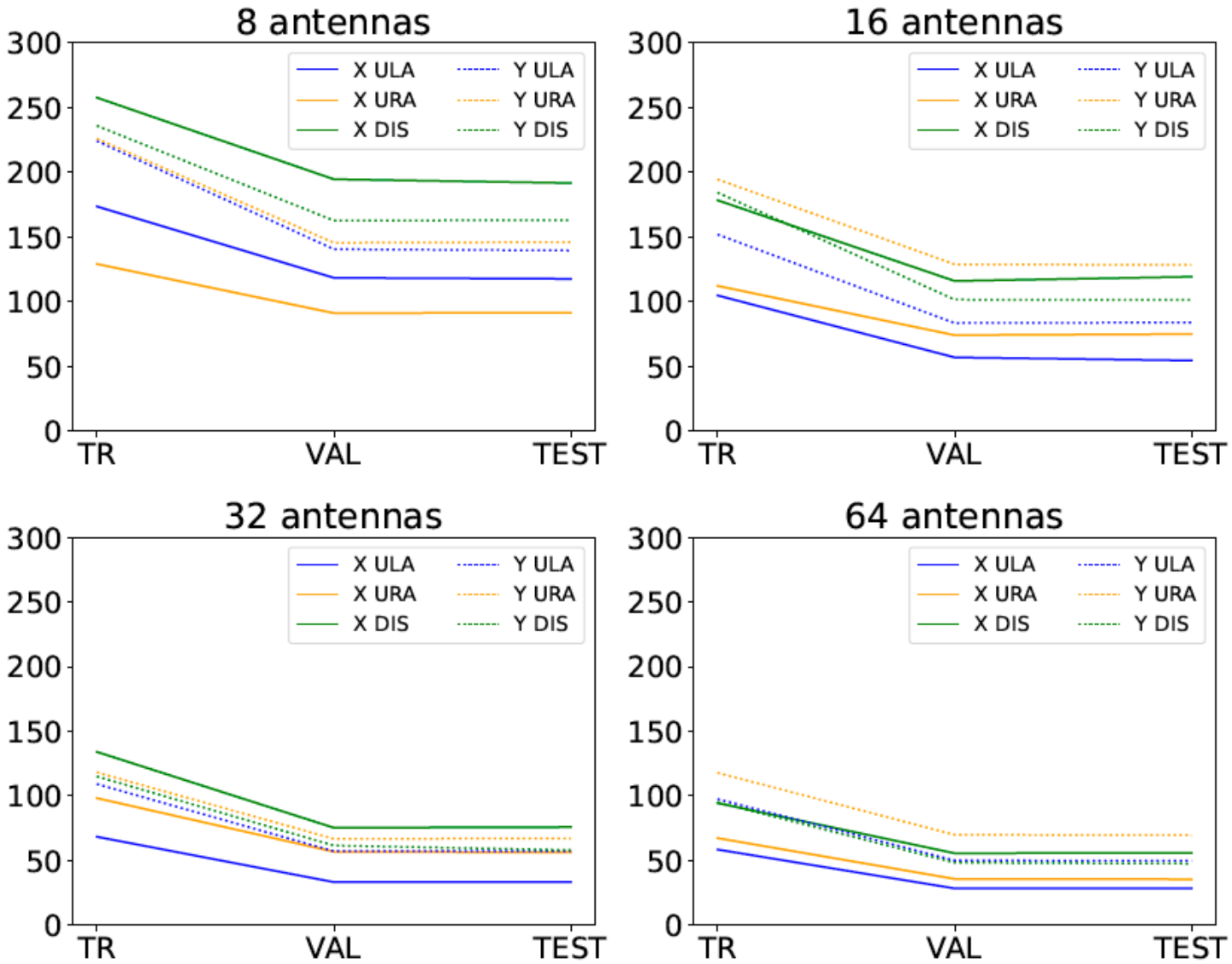
Table 2. RMSE (in mm) in validation (Val) and test split. Best results are shown in bold.

Algorithm	PosX		PosY	
	Val	Test	Val	Test
BR	226.05	225.00	251.43	255.54
ET	163.15	161.65	180.00	185.70
HGB	194.10	194.97	236.55	236.46
KNN	**110.50**	**110.54**	**133.70**	**140.16**
LiR	383.05	386.95	432.83	439.10
MLP	179.80	178.82	326.11	334.76
RF	226.09	225.18	251.37	255.62
RCV	383.04	386.94	432.80	439.06
XGB	178.41	180.03	202.45	201.66
LGB	194.14	194.15	231.19	232.89

Table 3. RMSE (in mm) for the different HyNNs architectures and HyViT in Validation (Val) and test. Best results are shown in bold.

Position	Model	TINTO		IGTD		REFINED	
		Val	Test	Val	Test	Val	Test
PosX	HyCNN	187.10	188.10	92.8	92.21	105.69	105.38
	HyTNN	178.28	179.25	119.59	119.62	115.90	114.98
	HyTTNN	181.96	184.19	179.01	180.05	193.56	196.09
	HyGTNN	176.71	176.43	173.42	174.20	173.38	174.02
	HyViT	**103.27**	**104.17**	**46.57**	**45.77**	**41.38**	**41.84**
PosY	HyCNN	152.19	151.94	101.01	99.45	115.40	114.69
	HyTNN	143.10	143.29	95.95	95.83	112.27	112.02
	HyTTNN	151.35	151.97	155.35	154.12	147.22	146.01
	HyGTNN	155.06	153.40	154.68	154.50	157.10	155.39
	HyViT	**121.77**	**123.90**	**70.84**	**68.93**	**90.11**	**90.56**

Inferencia


8 antennas
16 antennas
32 antennas
64 antennas
X ULA
X URA
X DIS
Y ULA
Y URA
Y DIS
300
250
200
150
100
50
0
TR
VAL
TEST

Trabajo futuro

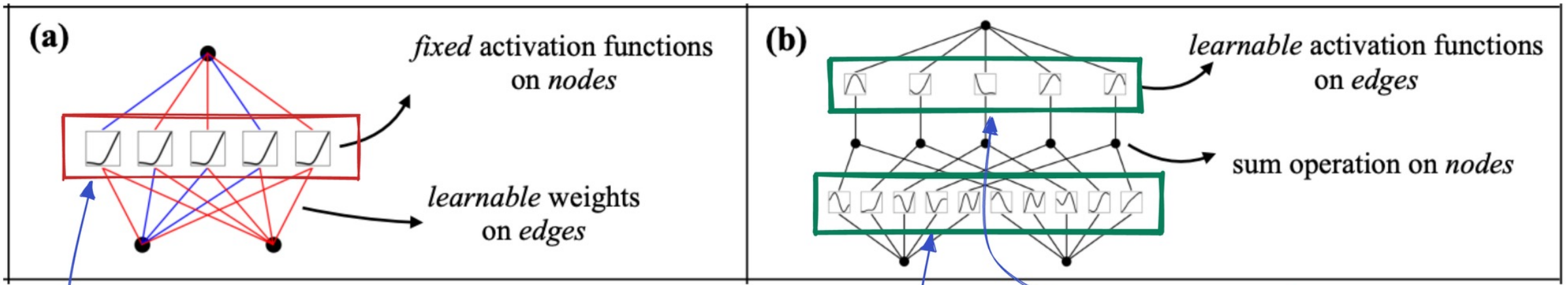
ETS de Ingeniería Informática

UNED

KAN: Kolmogorov-Arnold Networks

Neural Network

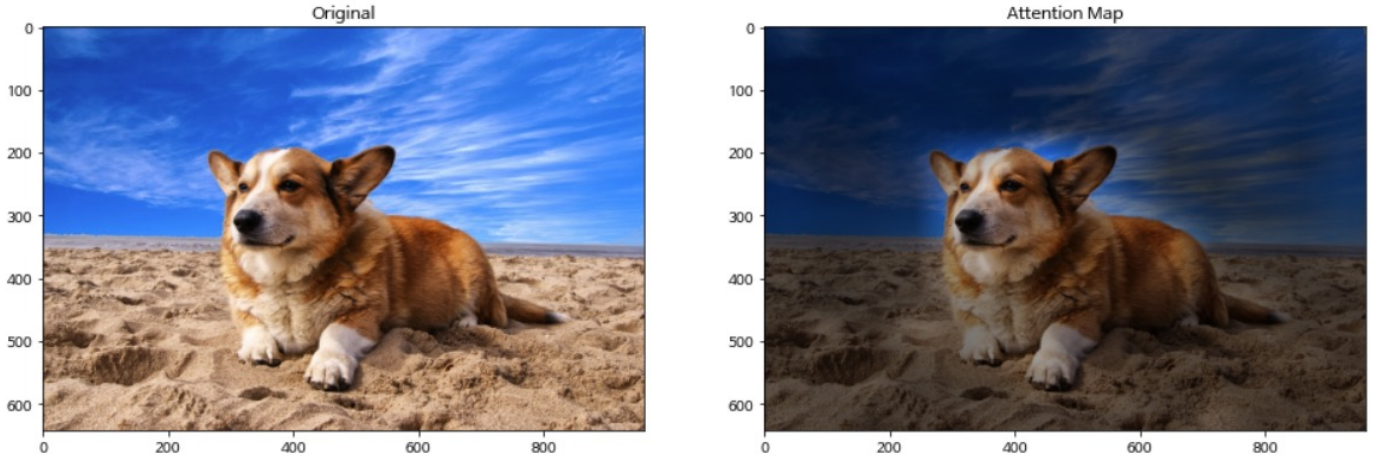
KAN

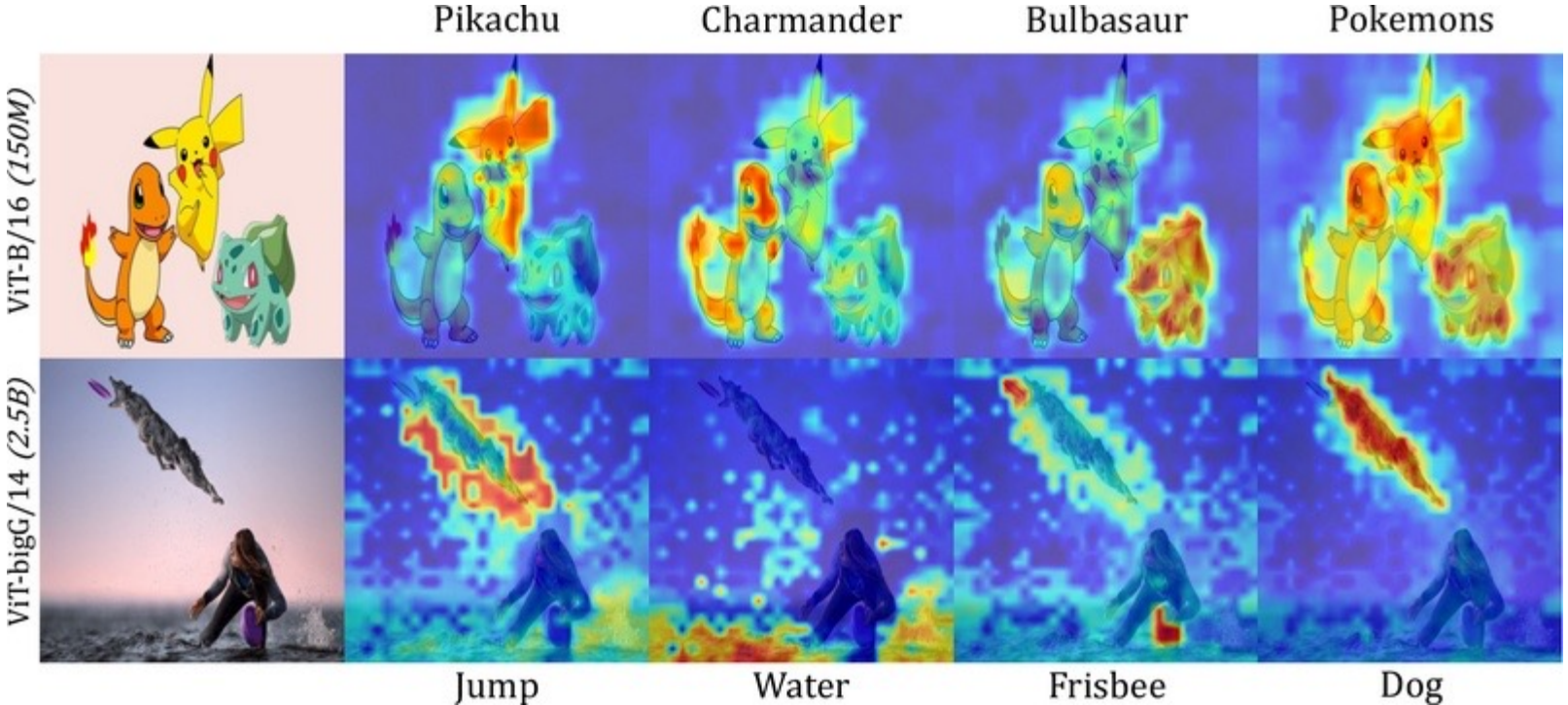


All neurons have a fixed activation function

Every edge has a different activation function

Explicabilidad





Bousselham, W., et al. (2024). LeGrad: An Explainability Method for Vision Transformers via Feature Formation Sensitivity. *arXiv preprint arXiv:2404.03214*.

Más información

ETS de Ingeniería Informática

UNED

Más información

- Documentación oficial de TINTOlib: https://tintolib.readthedocs.io/en/latest/
- TINTOlib Crash Course: https://github.com/oeg-upm/TINTOlib-Crash_Course
- Librería TINTOlib en PyPI: https://pypi.org/project/TINTOlib/
- GitHub con el código de TINTOlib: https://github.com/oeg-upm/TINTOlib
- GitHub con el código de TINTO: https://github.com/oeg-upm/TINTO
- Artículo sobre TINTO y su aplicación en indoor localization. Incluye la definición formal matemática: https://doi.org/10.1016/j.inffus.2022.10.011
- Artículo sobre TINTO: https://doi.org/10.1016/j.softx.2023.101391

¡Gracias!

Dr. Manuel Castillo-Cara
www.manuelcastillo.eu
Departamento de Inteligencia Artificial
Escuela Técnica Superior de Ingeniería Informática
Universidad Nacional de Educación a Distancia (UNED)