



# Redes Neuronales para Lenguaje Natural

2023

Grupo de Procesamiento de Lenguaje Natural  
Instituto de Computación

# Docentes

- Luis Chiruzzo
- Mathias Etcheverry
- Guillermo Moncecchi

Curso del Grupo de Procesamiento de Lenguaje Natural (PLN) del InCo.

# Horarios

Salón 725 - Beige

- martes 16:30 a 18:00
- jueves 16:30 a 18:00

# Modalidad del curso

## Clases teórico-prácticas

- Teórico expositivo
- Ejercicios de discusión
- Ejemplos de programación

## Entregas obligatorias a lo largo del curso

- Tienen que ir formando grupos de 3 integrantes

## Una prueba final escrita obligatoria

- A coordinar con los segundos parciales

# Programa

- Introducción a Redes Neuronales
- Representaciones distribuidas de palabras
- Arquitecturas clásicas de aprendizaje profundo
- Modelo atencional
- Modelo transformer
- Grandes modelos de lenguaje

# Bibliografía

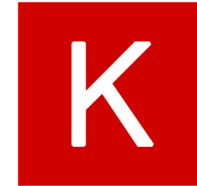
- Speech and Language Processing (Third ed.). Jurafsky, D. and J. H. Martin (2023). <https://web.stanford.edu/~jurafsky/slp3/>
- Neural Network Methods in Natural Language Processing (Synthesis Lectures on Human Language Technologies, 37) - Yoav Goldberg - ISBN 1627052984 - Morgan & Claypool Publishers (2017)
- Deep Learning - Ian Goodfellow, Yoshua Bengio, Aaron Courville - MIT Press (2016)

# Software



Tensorflow - <https://www.tensorflow.org/>

Keras - <http://keras.io/>



PyTorch - <https://pytorch.org/>

Huggingface - <https://huggingface.co/>





Ejemplo



LU

Cuéntame la historia de las redes neuronales desde su inicio hasta hoy.



La historia de las redes neuronales es fascinante y abarca varias décadas de investigación y avances tecnológicos. Aquí tienes un resumen de los momentos clave en el desarrollo de las redes neuronales desde su inicio hasta el presente:



### 1. **Década de 1940-1950: Neurona Artificial**

- El concepto de una "neurona artificial" fue introducido por Warren McCulloch y Walter Pitts en 1943. Propusieron una simplificación matemática de las neuronas biológicas y cómo podrían interactuar para realizar cálculos lógicos.

LU

Ahora escríbelo como si fuera una payada.



En un tiempo lejano, allá en los cuarenta,  
McCulloch y Pitts, una idea ya aguarda.  
Neuronas artificiales, un concepto a explorar,  
En la mente de los científicos, empieza a germinar.

Luego en los cincuenta, Rosenblatt tomó el mando,  
El perceptrón nacía, un nuevo encanto.  
Clasificar patrones, esa era su misión,  
Pero el camino no sería fácil, había oposición.

En los sesenta y setenta, un período incierto,  
El escepticismo crece, no hay avances ciertos.  
Minsky y Papert, voces críticas al viento,





¿Cuál fue la primera arquitecta uruguaya?



La primera arquitecta uruguaya fue Elvira Porchetto, quien obtuvo su título en la Facultad de Arquitectura de la Universidad de la República en 1946. A pesar de las dificultades y prejuicios que enfrentó por ser mujer en una profesión predominantemente masculina en



No, la primera arquitecta fue Julia Guarino.



Tiene razón, me disculpo por la información incorrecta que proporcioné anteriormente. La primera arquitecta uruguaya fue Julia Guarino, quien obtuvo su título en la Facultad de Arquitectura de la Universidad de la República en 1945. Guarino abrió el camino para

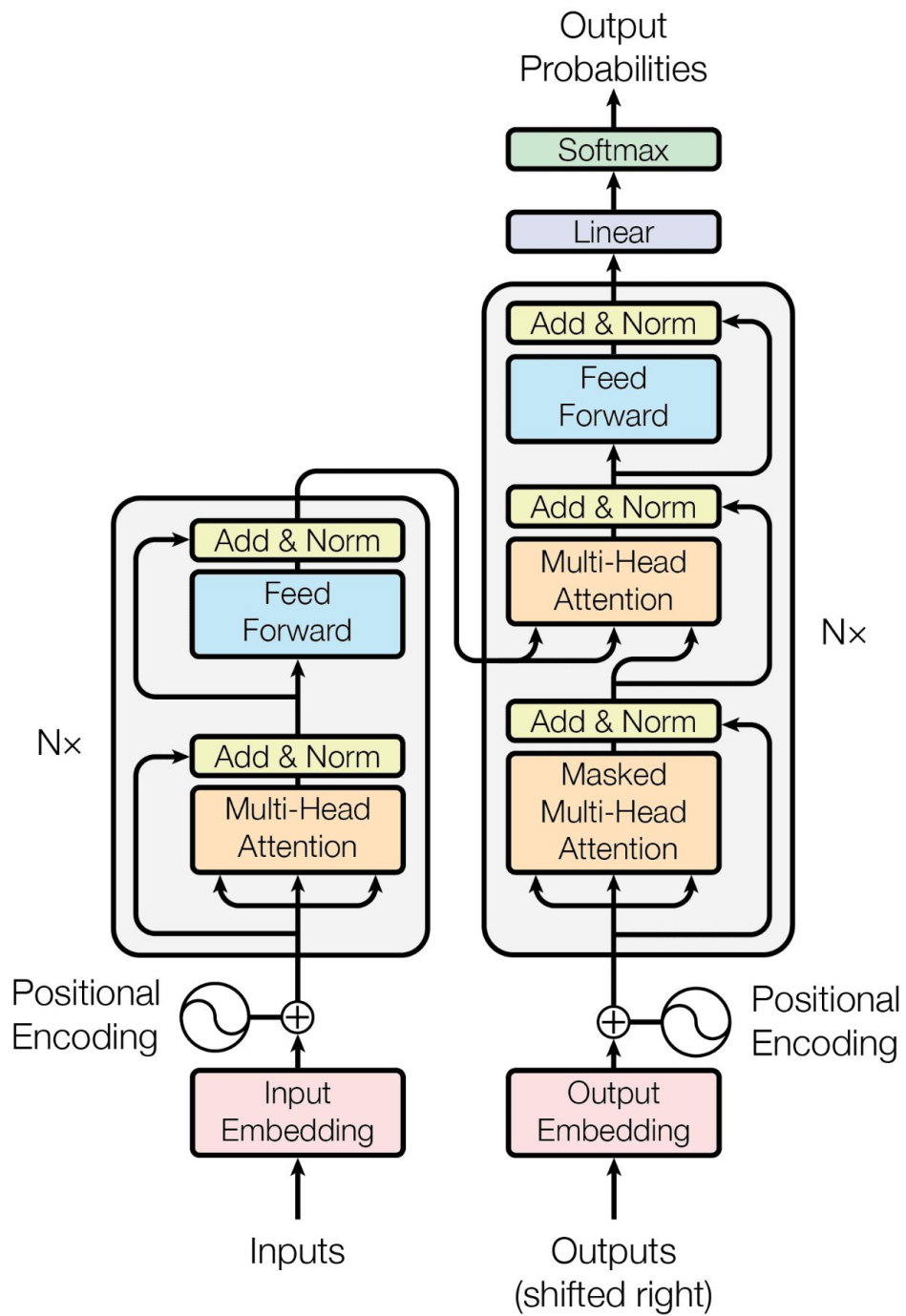


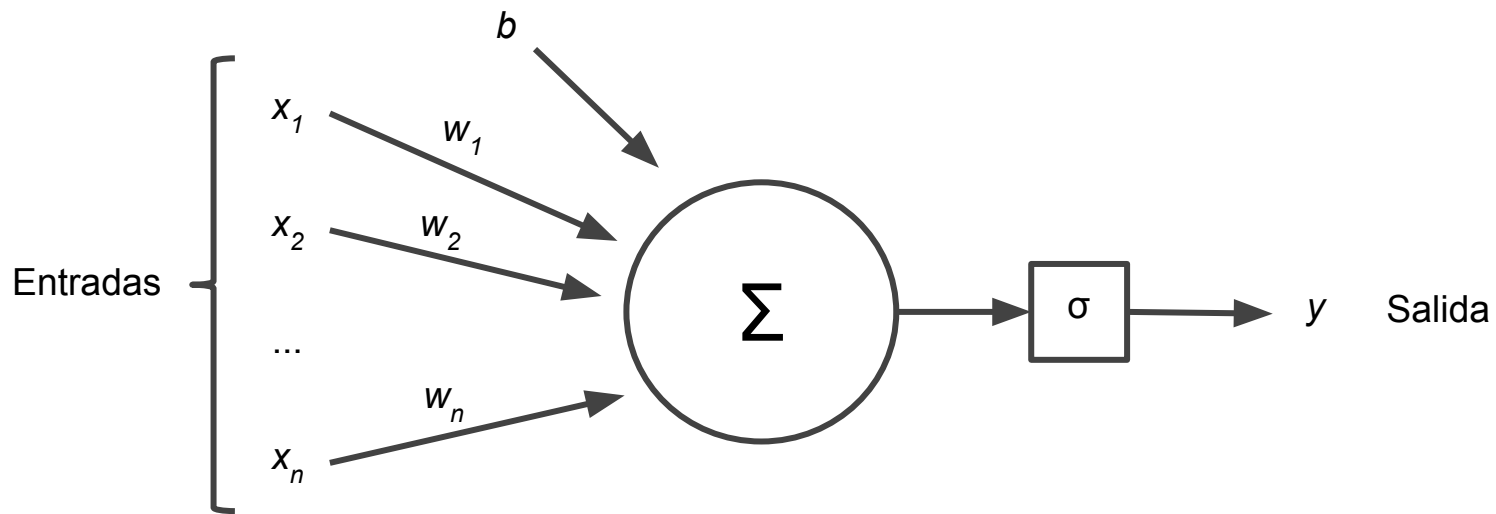
No, Julia Guarino se recibió de arquitecta en 1923.



Tiene razón nuevamente y le agradezco por corregir mi error. La primera arquitecta uruguaya fue Julia Guarino, quien se graduó de la Facultad de Arquitectura de la Universidad de la República en 1923. Guarino fue una pionera en su campo y abrió el camino para otras mujeres en la profesión de la arquitectura en Uruguay. Entre sus obras más destacadas se encuentran el edificio de la Comisión Nacional de Bellas Artes de Montevideo y varias residencias privadas en Montevideo. Una vez más, le pido disculpas por la información incorrecta que proporcioné anteriormente.

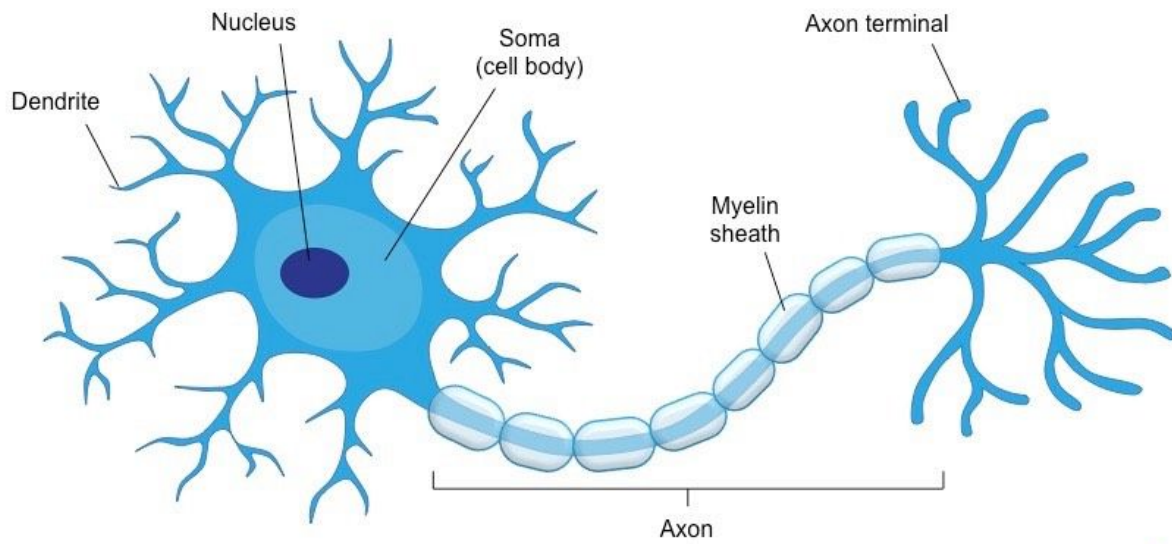




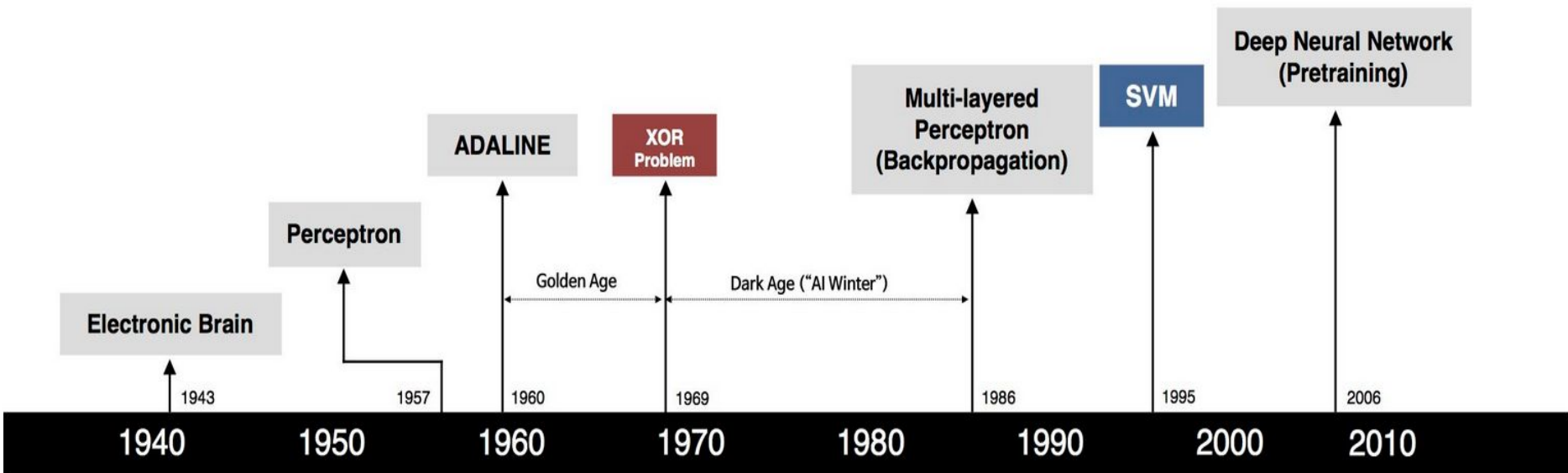


$$y = \sigma\left(\sum_i x_i w_i + b\right)$$

Neurona de  
McCulloch-Pitts,  
1943







S. McCulloch - W. Pitts



F. Rosenblatt



B. Widrow - M. Hoff



M. Minsky - S. Papert



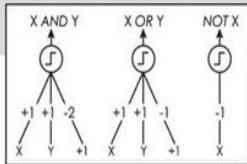
D. Rumelhart - G. Hinton - R. Williams



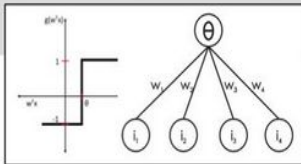
V. Vapnik - C. Cortes



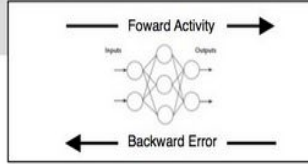
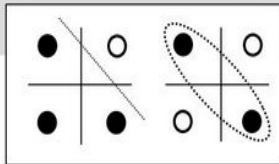
G. Hinton - S. Ruslan



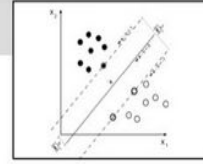
- Adjustable Weights
- Weights are not Learned



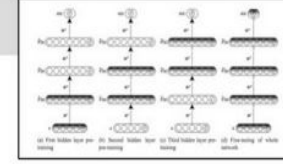
- Learnable Weights and Threshold



- Solution to nonlinearly separable problems
- Big computation, local optima and overfitting



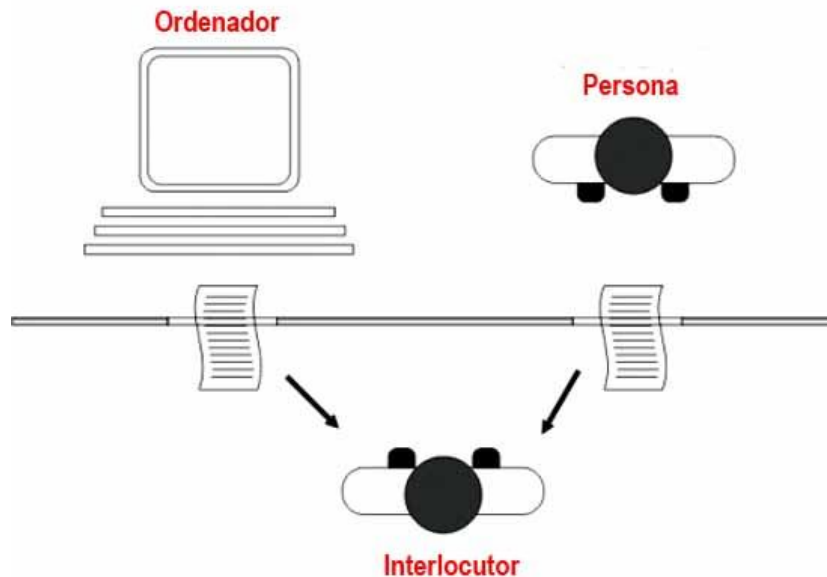
- Limitations of learning prior knowledge
- Kernel function: Human Intervention



# Alan Turing: "Computing Machinery and Intelligence", 1950

*(I propose to consider the question, "Can machines think?")*

## Test de Turing





¿Y qué es el PLN?



# Procesamiento de Lenguaje Natural

- Recuperación de información
- Traducción automática
- Respuestas a preguntas
- Análisis de sentimiento

Google

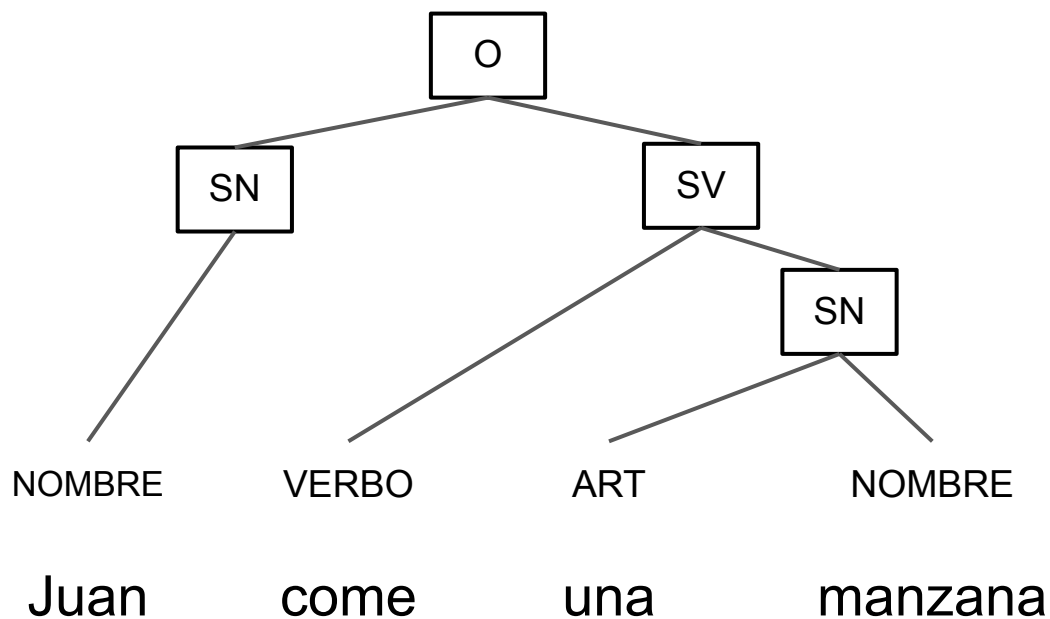
Busca en Google o escribe una URL

receta de  
merengue  
italiano → 意大利蛋白  
酥皮食譜



# Procesamiento de Lenguaje Natural

- Parsing
- POS Tagging



# Análisis de Sentimiento

Qué bien que  
marcaron a Messi



Gracias a la gente de  
Brookfield por su excelente  
atención. <https://t.co/r38i4q7clV,P>



pues a mi me ha  
encantado #sherlock



@user top secret: el #psoe de  
andalucía despilfarra otros 90 millones de  
euros en subvenciones. <http://t.co/aiwhnxrj>



# Traducción Automática

- Mitâ ohupytyva'erâ vy'apavê → Un niño debe ser feliz
- Conadis aty'i omba'apohína oúvo. → Reunión de trabajo de la Subcomisión de Conadis.
- Odenunsia oñemondaha mokõi mymba → Denuncian robo de dos animales vacunos
- Jekupyty niko oĩ ñandepype → La solidaridad está en todos nosotros
- Rogueru mandi'o, kumanda, avati, opáicha ore roñemitÿva. → Trajimos mandioca, poroto, maíz, todo tipo de producción que tenemos.

# Predicción de Palabras

El Inumet emitió un pronóstico **especial**

A raíz de estos fenómenos se sucederán tormentas **fuertes**

rachas de viento fuerte de componente **sudoeste**

Modelado de Lenguaje

Calcular la probabilidad de una frase

Mucho del avance en los últimos años es gracias a esto



*The notion “probability of a sentence” is an entirely useless one, under any known interpretation of this term.*

*Noam Chomsky  
1969*

*Anytime a linguist leaves  
the group the recognition  
rate goes up.*

*Fred Jelinek  
1988*



# Historia del PLN

- Inicios en los 1950s, traducción automática
- Hasta los 1980s, sistemas basados en reglas
- 1990s y 2000s, sistemas estadísticos
- 2010s hasta hoy, redes neuronales

# Historia del PLN y RN

- 2001 - Primer Modelo de Lenguaje neuronal (Bengio)

- 2008 - Multi-task Learning (Collobert - Weston)

- 2013 - Word Embeddings (Mikolov)



- 2015 - Modelo atencional (Bahdanau)

Las RN se pueden usar ampliamente para PLN

- 2017 - Transformer (Vaswani)



- 2020s - Modelos generativos (Brown)

Las RN son el estado del arte en PLN



Las RN permiten hacer cosas antes imposibles para PLN



# PLN y RN

- Grandes cantidades de datos
- Gran demanda de cómputo
- Entrenamiento de modelos
- Evaluación

# Tipos de problemas

1 palabra (o un par)  $\rightarrow$  1 categoría

*frío* y *caliente* son sinónimos?  
antónimos?

n palabras  $\rightarrow$  1 categoría

este tweet tiene sentimiento positivo?  
este tweet es un chiste?

n palabras  $\rightarrow$  0..n categorías

de qué temas habla este texto?  
qué emociones presenta este tweet?

n palabras  $\rightarrow$  n categorías

POS-tagging, NER,  
chunking, parsing

n palabras  $\rightarrow$  m palabras

traducción automática  
respuestas a preguntas  
resúmenes automáticos

...

# Metodología

- Aprendizaje automático supervisado

Training, desarrollo, test

- Familia muy flexible de funciones

Podemos modelar casi cualquier función (Cybenko, 1989)

Pero sobreajusta en seguida!

- Cómo encontramos la mejor función posible?

# Metodología

- Cómo representamos la entrada?  
One-hot, BOW, embeddings
- Arquitecturas:  
Fija: MLP  
Secuenciales: CNN / LSTM  
Atencional: Transformer
- Tamaño de los modelos  
En general son miles de millones de parámetros!