

Curso: Aprendizaje Supervisado



Dr. Edgar Eloy Carpio Vargas
Universidad Nacional del Altiplano
Puno-Perú


SILABO

FACULTAD
ESCUELA PROFESIONAL
PROGRAMA DE ESTUDIOS

INGENIERIA ESTADISTICA E INFORMATICA
INGENIERIA ESTADISTICA E INFORMATICA
CARRERA PURA

I. INFORMACIÓN GENERAL

I.1 Identificación Académica

a) Curso		APRENDIZAJE SUPERVISADO
b) Código		EST320
c) Prerequisito		EST314 - APRENDIZAJE NO SUPERVISADO
d) Número de Horas		03h teóricas, 02h prácticas, 04h virtuales, Total 05 horas
e) Créditos		04
f) Número de Horas virtuales		04
g) Año y Semestre Académico		2024-II
h) Ciclo de Estudios		VII
i) Duración		Del 19 de Agosto al 20 de Diciembre del 2024 (18 semanas)
j) Área Curricular		Estudios de especialidad
k) Características del Curso		teórico practico

I.2 Docente

a) Apellidos y Nombres	CARPIO VARGAS EDGAR ELOY
b) Condición y Categoría	nombrado - principal
c) Especialidad	ING.ESTADISTICO
a) Apellidos y Nombres	PAREDES QUISPE JUAN REYNALDO
b) Condición y Categoría	nombrado - principal
c) Especialidad	ING.ESTADISTICO

I.3 Ambiente donde se realizó el aprendizaje

- a) 203- laboratorio 1

VI. TRATAMIENTO DE UNIDADES DIDÁCTICAS

UNIDAD 1		UNIDAD 1
LOGROS DE APRENDIZAJE DE LA UNIDAD Construye modelos de regresión utilizando técnicas learning.		
TIEMPO DE DESARROLLO		Del 19 de Agosto al 21 de Octubre del 2024 (Total 45 horas)
HORAS DE ENSEÑANZA VIRTUAL/UNIDAD		02
SEMANAS	CRITERIOS DE DESEMPEÑO	CONOCIMIENTOS
Semana 1	retroalimentacion de saberes previos	revisión de técnicas no supervisadas. factorial, cluster y otros). Limpieza de datos
Semana 2	Resuelve casos de estudio relacionados a modelo lineal simple y multiple	- Regresión lineal simple - Regresión lineal multiple - Regresión lineal utilizando técnicas machine learning
Semana 3	Mejora procedimientos del uso de regresión lineal simple y multiple	Retroalimentación
Semana 4	Resuelve casos de estudio relacionados a arboles de regresión	- Arboles de regresión - Evaluación y predicción de arboles de regresión
Semana 5	Mejora procedimientos de uso de arboles de regresión	Retroalimentación
Semana 6	Resuelve casos de estudio relacionados a maquinas de soporte de vectores	- Máquinas de soporte de vectores - evaluación y predicción
Semana 7	Mejora procedimientos de uso de maquinas de soporte de vectores	Retroalimentación
Semana 8	Resuelve casos de estudio relacionados a redes neuronales	Introducción a redes neuronales
Semana 9	Mejora los procedimientos de uso de modelos de regresión	Retroalimentación y evaluación
PORCENTAJE DE AVANCE ACADÉMICO DE LA UNIDAD: 50%		

UNIDAD 2		UNIDAD 2	
LOGROS DE APRENDIZAJE DE LA UNIDAD			
Las técnicas de aprendizaje learning son usadas para obtener modelos estadísticos de clasificación			
TIEMPO DE DESARROLLO		Del 21 de Octubre al 20 de Diciembre del 2024 (Total 45 horas)	
HORAS DE ENSEÑANZA VIRTUAL/UNIDAD		02	
SEMANAS	CRITERIOS DE DESEMPEÑO		CONOCIMIENTOS
Semana 10	Resuelve casos de estudio relacionados a modelo discriminante		Análisis discriminante lineal y cuadrático
Semana 11	Mejora procedimientos de uso de modelos discriminante		Retroalimentacion
Semana 12	Resuelve casos de estudio relacionados a modelo de regresión logística simple y múltiple		Regresión logística simple y múltiple
Semana 13	Mejora procedimientos del uso de regresión logística		Retroalimentacion
Semana 14	Resuelve casos de estudio relacionados a arboles de clasificación		Arboles de clasificación
Semana 15	Mejora procedimientos del uso de regresión logística		Retroalimentación
Semana 16	Resuelve casos de estudio relacionados al uso de maquina de soporte de vectores		Máquina de soporte de vectores
Semana 17	Mejora procedimientos del uso MSV		Retroalimentacion
Semana 18	Resuelve casos de estudio relacionados a modelos de clasificacion Evaluacion final		_ otros modelos de clasificacion (Boosting, Baggin) - evaluacion y retroalimentacion
PORCENTAJE DE AVANCE ACADÉMICO DE LA UNIDAD: 50%			

IX. PRODUCTO DE APRENDIZAJE

FECHA DE PRESENTACIÓN	PRODUCTO
ultima semana	Portafolio

X. EVALUACIÓN DE APRENDIZAJE

X.1 Evidencias, indicaciones, técnicas e instrumentos de evaluación

UNIDAD	LOGROS DE APRENDIZAJE	EVIDENCIAS DESEMPEÑO: De acción, objeto o producto (%)	PONDERACIÓN (Obligatorio en base a 100%)	TÉCNICAS	INSTRUMENTOS
1	Construye modelos de regresión utilizando técnicas learning.	- Reporte de hallazgos de casos de estudio - presentación de portafolio - Retroalimentacion	50%	- Observacion	- Rubrica
2	Las técnicas de aprendizaje learning son usadas para obtener modelos estadísticos de clasificación	- Reporte de presentación de casos de estudio - portafolio - retroalimentación	50%	Observacion	rubrica

X.2 Calificación:

La fórmula para la obtención del promedio final del curso es la siguiente:

$$\text{Promedio Final} = (50\%)IUPP + (50\%)IIUPP$$

Donde:

IUPP : Primero unidad promedio parcial

IIUPP : Segundo unidad promedio parcial

promedio de primera unidad = $0.50(LA1) + 0.30(LA2) + 0.20(LA3)$

LA1: logro de aprendizaje 1

LA2: logro de aprendizaje 2

LA3: logro de aprendizaje 3

Inteligencia Artificial-Introducción

Breve reseña histórica



. LIX. No. 236.]

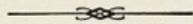
[October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY



I.—COMPUTING MACHINERY AND INTELLIGENCE

By A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to

LA IA Tiene un periodo de estudio desde 1950

A.M. Turing

¿Pueden la maquinas pensar?

Limitaciones:

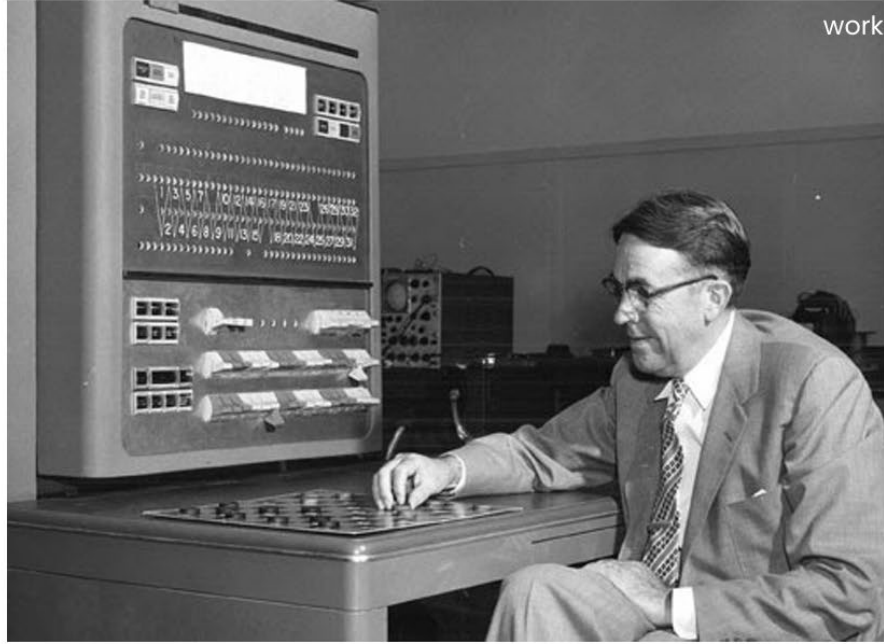
en 1950 existían los tubos de vacío (televisores) no existían los transmisores, entonces no existía la capacidad computacional.

¿Cómo puedo estar seguro, si con quien estoy hablando es una persona o una computadora?



1956 Artificial intelligence

Se desarrollo el lenguaje LISP que permitió desarrollar algunos modelos y se acuña el termino
“Inteligencia artificial”



Checkers

1952 Arthur L. Samuel

Arthur L. Samuel desarrolló un juego llamado Checkers, un juego similar a las damas, en el cual se podía jugar con la computadora y tenía cierta capacidad de aprendizaje (con tubos de vacío)

Por 1950 se vivió el mismo boom que vivimos ahora con la inteligencia artificial, pero se toparon con limitaciones computacionales.

Ya por 1959 se desarrolla en transmisor.

Machines will be capable, within twenty years, of doing any work a man can do. —Herbert Simon

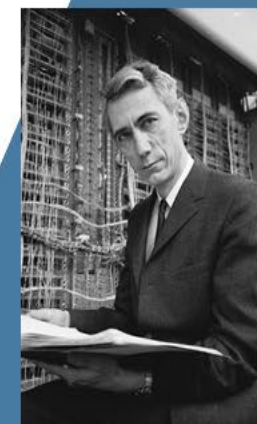
Decían, en 20 años la computadora va a tener la misma capacidad de los humanos, lo mismo que debatimos ahora.

Within 10 years the problems of artificial intelligence will be substantially solved. —Marvin Minsky

Dentro de 10 años, la IA va a solucionar muchos problemas, lo mismo que estamos pensando ahora (es como retrocederíamos a las mismas preguntas).

I visualize a time when we will be to robots what dogs are to humans, and I'm rooting for the machines. —Claude Shannon

Visualiza e indica que nosotros vamos a ser para los robots como los perros son para los humanos



En 1966 cuando ya habían transmisores, surgió la pregunta ¿cómo aplicamos la inteligencia artificial al lenguaje (aparece pascal, se hace regresión, clasificación etc). Veamos el siguiente ejemplo de traducción, no es una traducción correcta, empeora cuando se quiere volver al lenguaje inicial. Entonces, se puso en duda la IA. Entonces se inició el **primer invierno de la IA**

The spirit is willing but the flesh is weak.



(Russian)



The vodka is good but the meat is rotten.



The ALPAC Report

The ALPAC (Automatic Language Processing Advisory Committee) was a govt. committee of seven scientists.

Their 1966 report was very skeptical of the progress in computational linguistics and machine translation.

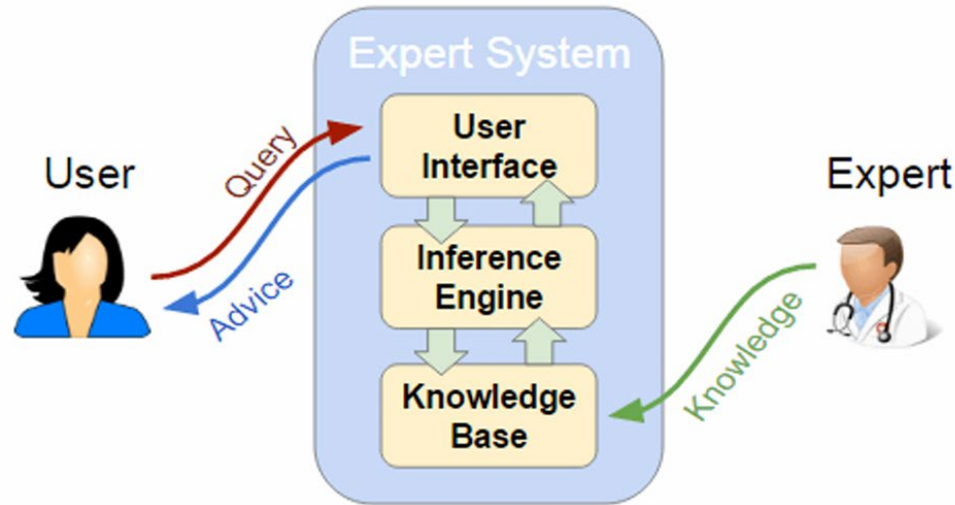
Inicio del primer invierno de la IA

¿Por qué se detuvo las investigaciones sobre IA?, ¿cuales fueron los problemas?

- Computación limitada
 - Hardware limitado
- Información limitada
 - Problemas complejos (Palabras, objetos, conceptos)

En los 70-80'

En estos años se trabajó los sistemas expertos,



Dominios específicos del conocimiento, de expertos, en forma de reglas

70 – 80's

- Las reglas se vuelven muy complejas de crear y mantener
- Las reglas no se adecuan ante lo incierto del mundo real

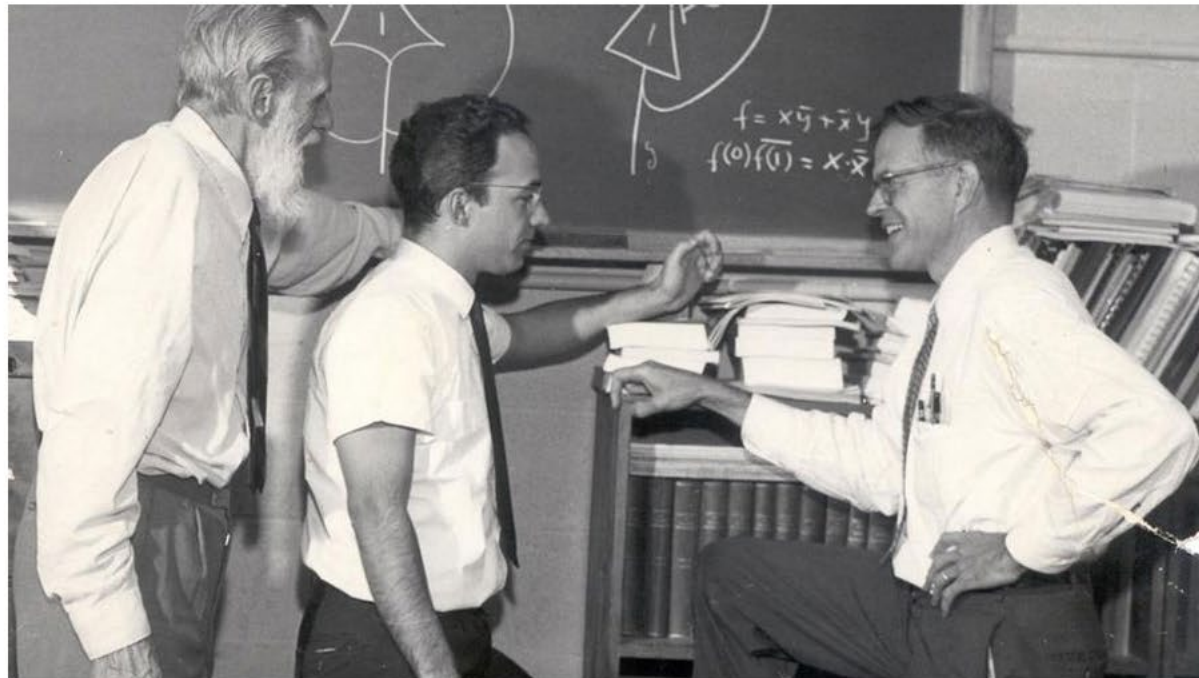
Paralelamente se desarrollaba otra área de desarrollo

Volvemos a 1943....

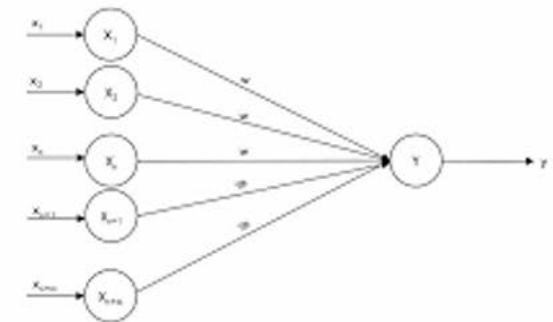
Redes neuronales...

Era una sola neurona

- 1943 La neurona de McCulloch-Pitts



McCULLOCH-PITTS NEURON: ARCHITECTURE (MP NEURON MODEL)



- 1958 Algoritmo Perceptron, para
clasificación lineal – Frank Rosenblatt

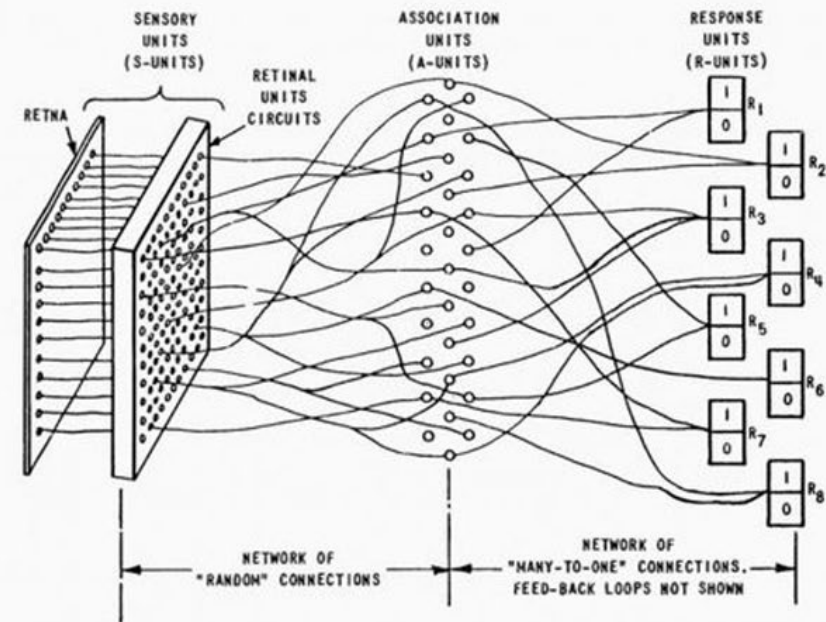
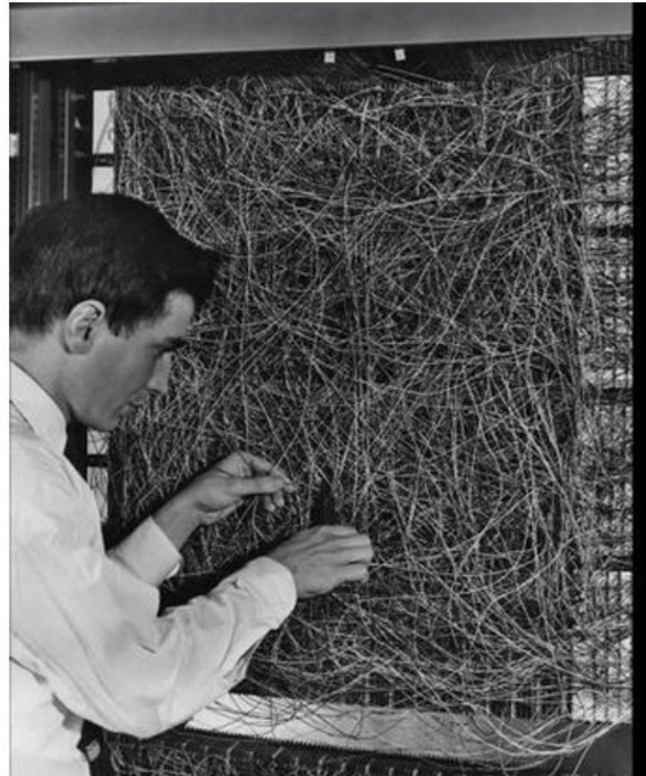
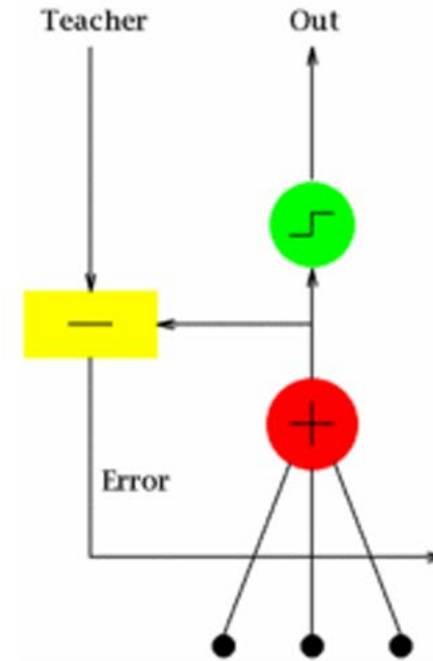
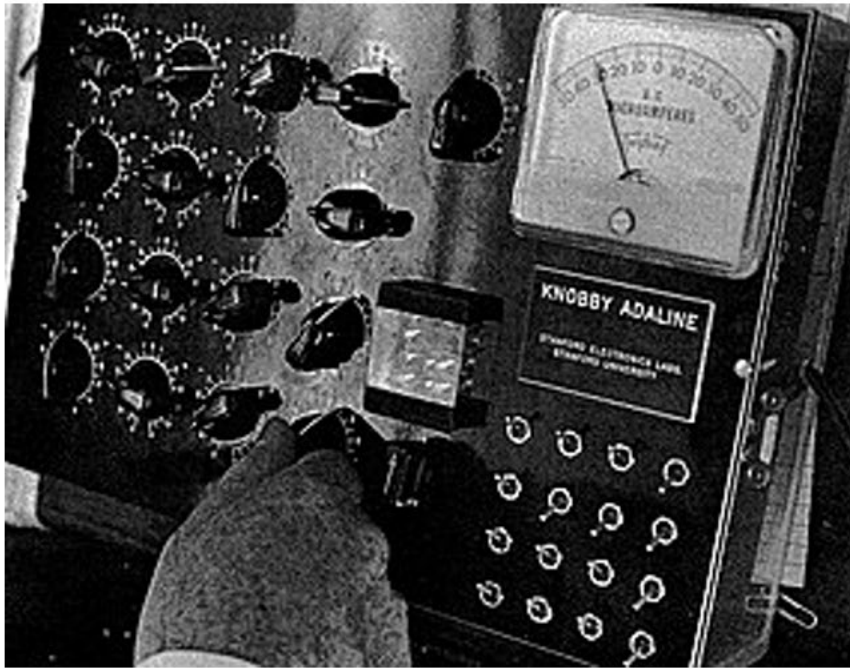


Figure 1 ORGANIZATION OF THE MARK I PERCEPTRON

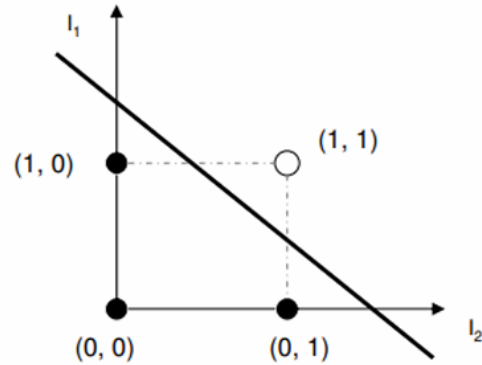
- 1959 ADALINE (Adaptive Linear Neuron), dispositivo para regresiones lineales



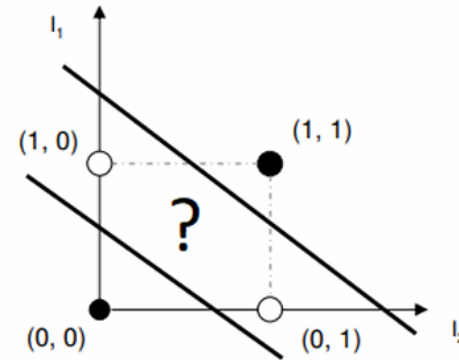
Los modelos perceptrón se utilizaba también para hacer funciones lógicas,

- 1969 Perceptron vs XOR

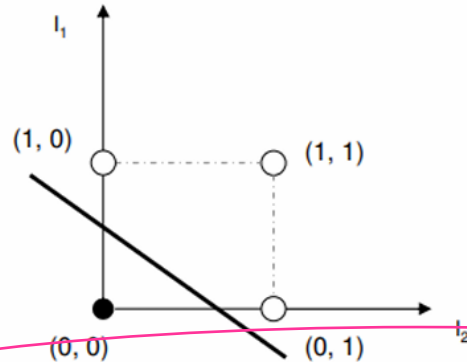
AND		
I_1	I_2	out
0	0	0
0	1	0
1	0	0
1	1	1



XOR		
I_1	I_2	out
0	0	0
0	1	1
1	0	1
1	1	0



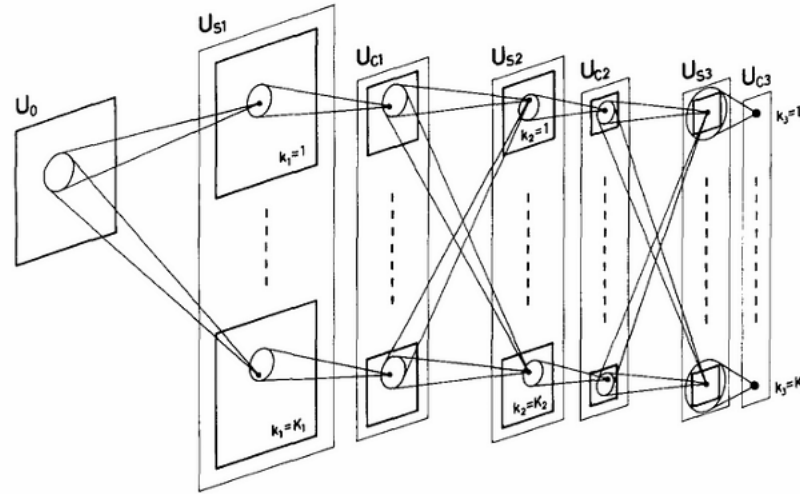
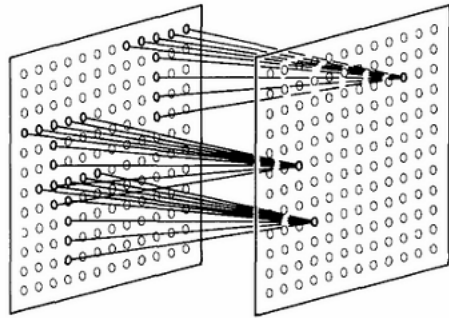
OR		
I_1	I_2	out
0	0	0
0	1	1
1	0	1
1	1	1



No se podía resolver el problema de XOR, porque solo se podía representar una neurona

Se detienen las investigaciones en redes neuronales

• 1980 Redes neuronales convolucionales



K. Fukushima

Para entonces:

- Ya se tenía computadoras
- Ya se podían manejar imágenes,
- Ya se podían procesar datos

Aparece los modelos convolucionales, que extraen características mas resaltante de una imagen

¿que se hace con esa información?, quedo pendiente

- 1986 Back propagation Rumelhardt, Hinton, Williams



- Desarrollaron la técnica back propagation,
- Entonces se soluciona el problema de que la XOR si se podía resolver con 3 neuronas. A eso se le llamo el perceptrón multicapa (redes neuronales).
- Aumenta la potencia computacional
- La Back propagation hace que las neuronas vayan aprendiendo (otra vez vuelve a crecer la IA)

- 1989 Redes neuronales convolucionales –
Lecun Es uno de pioneros de IA actuales



- Se tenia: back propagation, modelos convolucionales, redes neuronales.
- Propone las redes neuronales convolucionales. Que era unir los modelos convolucionales con todo lo desarrollado hasta entonces.
- Esta redes es la que estamos usando ahora.

Lecun, desarrolla el siguiente modelo. La computadora reconoce números escritos a mano



- Predecir los números
- Este modelo superaba a humanos. Reconocía números mejor que un humano
- Se inicio a usar dos dimensiones (imágenes)

En los 80. empezó otra corriente en paralelo a IA. Creación Internet y se inicio a generar datos

El modelo de Lecun genero otros modelos y se inicio el Deep Learning (solucionar modelos mas complejos)

- Deep learning

- 2012 Alexnet
- 2016 Alphago
- 2023 ChatGPT
- 2024 Llama, Claude

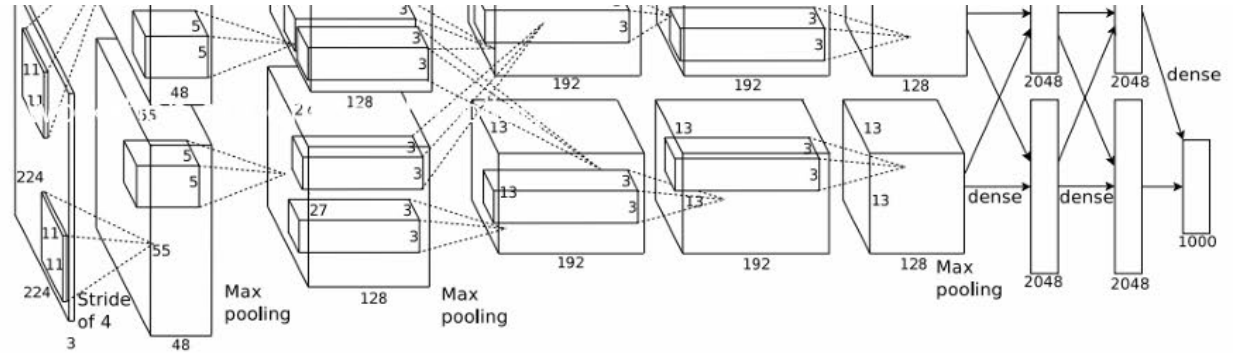
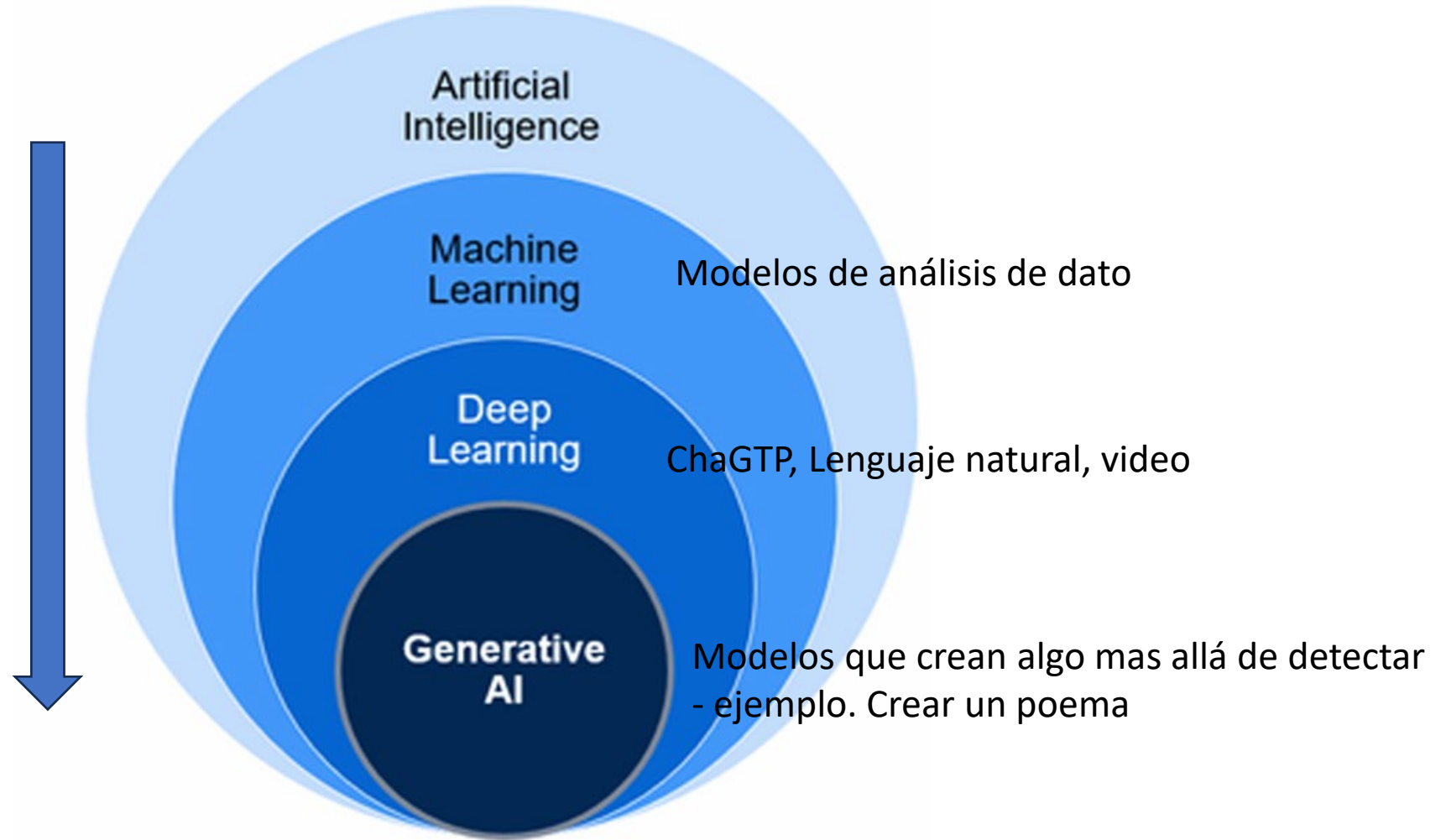


Figure 2: An illustration of the architecture of our CNN, explicitly showing the delineation of responsibilities between the two GPUs. One GPU runs the layer-parts at the top of the figure while the other runs the layer-parts at the bottom. The GPUs communicate only at certain layers. The network's input is 150,528-dimensional, and the number of neurons in the network's remaining layers is given by 253,440–186,624–64,896–64,896–43,264–4096–4096–1000.



- 2012 apareció otro problema. Cantidad de datos.
- un CPU es serializado
- Alexnet. Desarrollaron tarjetas graficas para juegos. Tienen miles de CPU que realizan cálculos rápidamente
- Procesamiento en paralelo para solucionar el problema
- Las empresas iniciaron a implementar sus centros de experimentación
- 2016. Alpha~~go~~. Es un juego mas complejo que ajedrez (el modelo aprendió a jugar)
- Se crearon muchos grupos de investigación
- 2023 chatGTP. Punto de inflexión.
- 2024 Llama, Claude, meta. Soluciones parecidas a ChatGTP. (ingenieros que salieron de open AI)

The AI Landscape



2 MACHINE LEARNING

Machine Learning (ML) es una subdisciplina de la inteligencia artificial que se centra en el desarrollo de algoritmos y modelos que permiten a las computadoras aprender y hacer predicciones o tomar decisiones basadas en datos.

Componentes

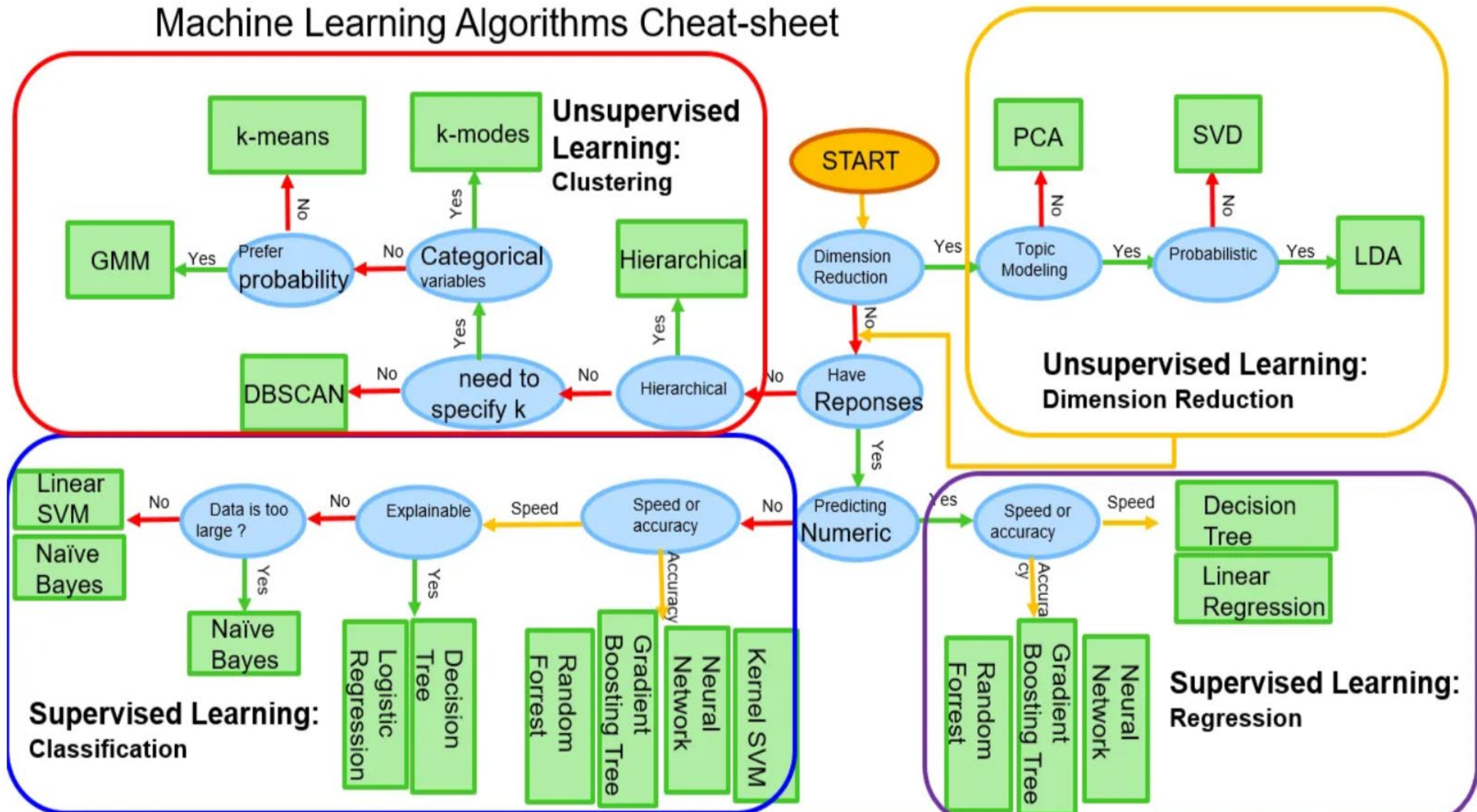
- Datos: Importancia de los datos en el entrenamiento de modelos.
- Modelos: Tipos de modelos (regresión, clasificación, etc.).
- Algoritmos: Ejemplos de algoritmos populares (regresión lineal, árboles de decisión, etc.).

Tipos de aprendizaje

- Aprendizaje supervisado
- Aprendizaje no supervisado.
- Aprendizaje por refuerzo.



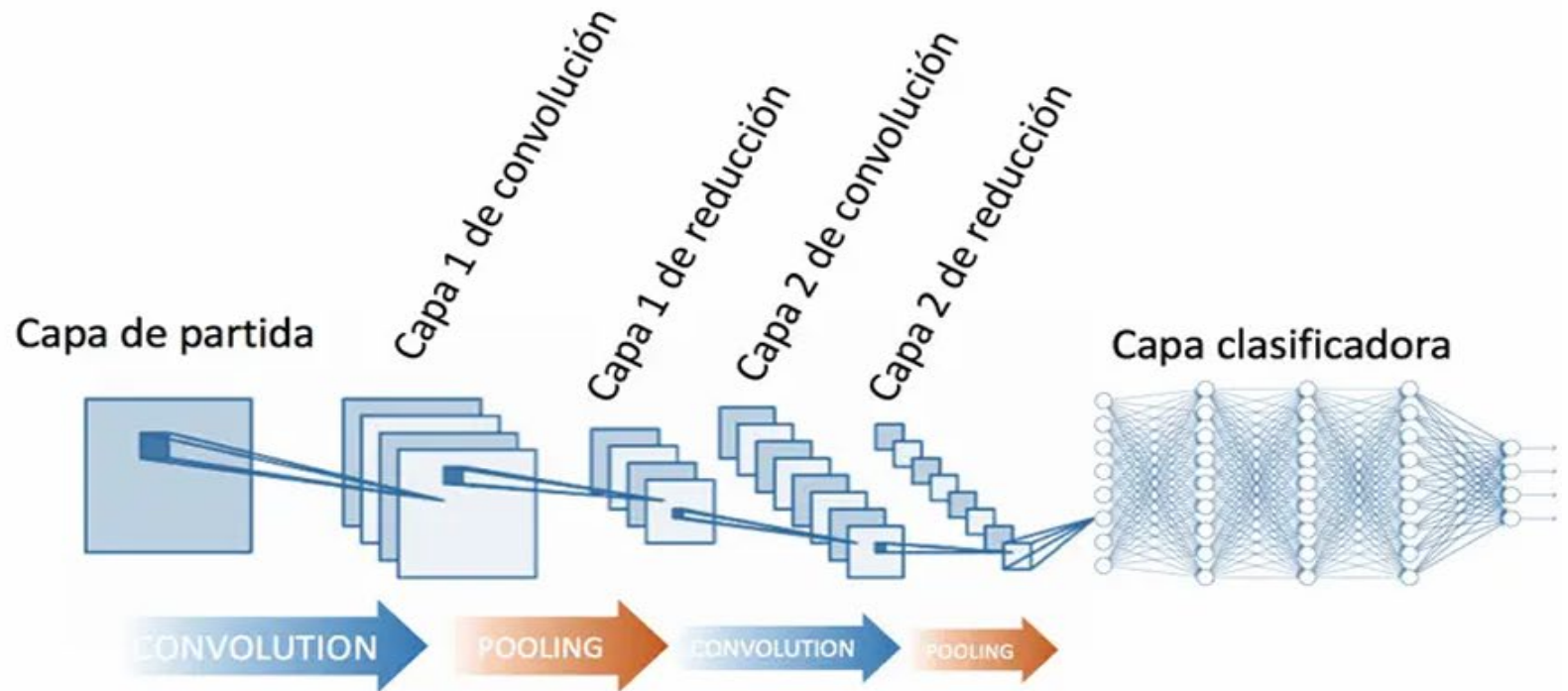
Machine Learning Algorithms Cheat-sheet



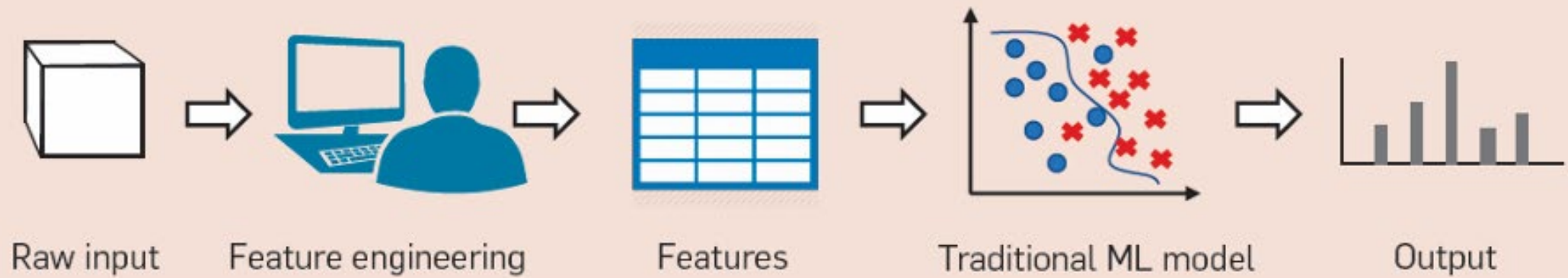
2 DEEP LEARNING

Componentes

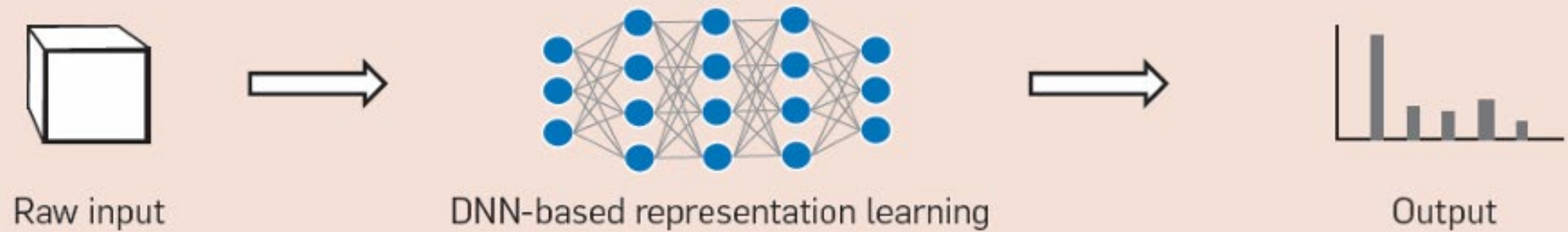
- Redes neuronales
- Redes neuronales convolucionales
- Transformers



Traditional machine learning



Deep learning



3 APLICACIONES

2024. empresa que están desarrollando modelos y solo nos dan para el uso



Meta Liberado. ver como funciona WhatsApp por ejemplo, y luego mejorar, etc. (800 GB de CPU)

4 ENTORNO DE DESARROLLO

Visual estudio



Colab diseñada por google



Frame works





gob.pe








Plataforma Nacional de Datos Abiertos

workshop at Dartmouth College

Datos Abiertos

Marco de Gobernanza de Datos del Estado Peruano está constituido por instrumentos técnicos y normativos que establecen los requisitos mínimos que las entidades de la Administración Pública deben implementar conforme a su contexto legal, tecnológico y estratégico para asegurar un nivel básico y aceptable para la recopilación, procesamiento, publicación, almacenamiento y apertura de los datos que administre.

Tipos de contenido

-  Recurso (10449)
-  Dataset (3658)
-  Entidades (246)
-  Harvest Source (27)
-  Página (3)
-  Data Dashboard (1)
-  Data Story (1)

Categorías

-  Economía y Finanzas (1235)
-  Gobernabilidad (1119)
-  Exprésate Perú con Datos (230)
-  Desarrollo Social (227)
-  Transporte (167)
-  Medio Ambiente y Recursos Naturales (139)
-  Salud (132)

14384 Distribución de Datos

Search

Ordenar por

Fecha cambiada

Pedido

Descendente

Consultar

Reiniciar



RENIEC - Consultas atendidas a la ciudadanía a través de los diferentes canales de comunicación [Registro Nacional de Identificación y Estado Civil]

 RENIEC - Registro Nacional de Identificación y Estado Civil

 Gobernabilidad

Cantidad de consultas atendidas a través de la Línea 0800, central telefónica, chat Institucional y correo institucional.

8x  



Reporte I Semestre 2024

 RENIEC - Registro Nacional de Identificación y Estado Civil

 RENIEC - Consultas atendidas a la ciudadanía a través de los diferentes canales de comunicación [Registro Nacional de Identificación y Estado Civil]





SIEN - Sistema de información del Estado Nutricional de niños y gestantes Perú - INS/CENAN