

MBA
USP
ESALQ

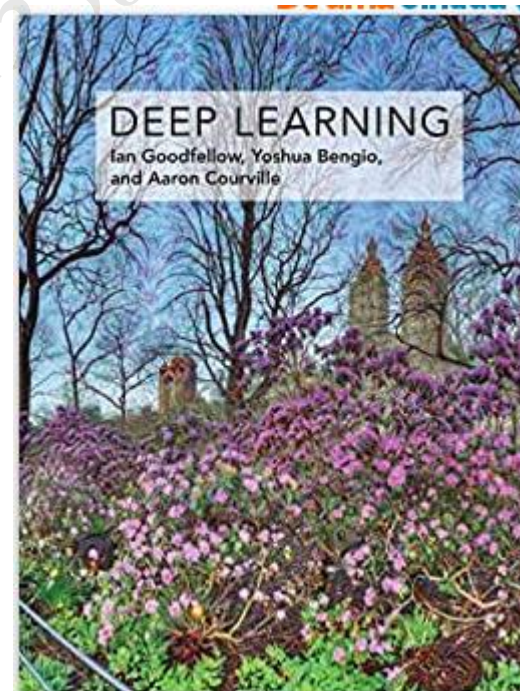
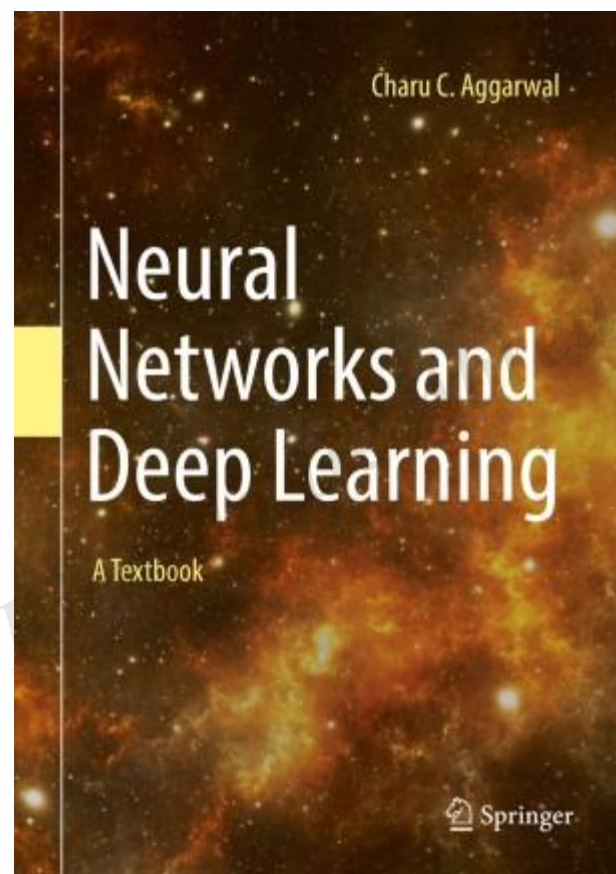
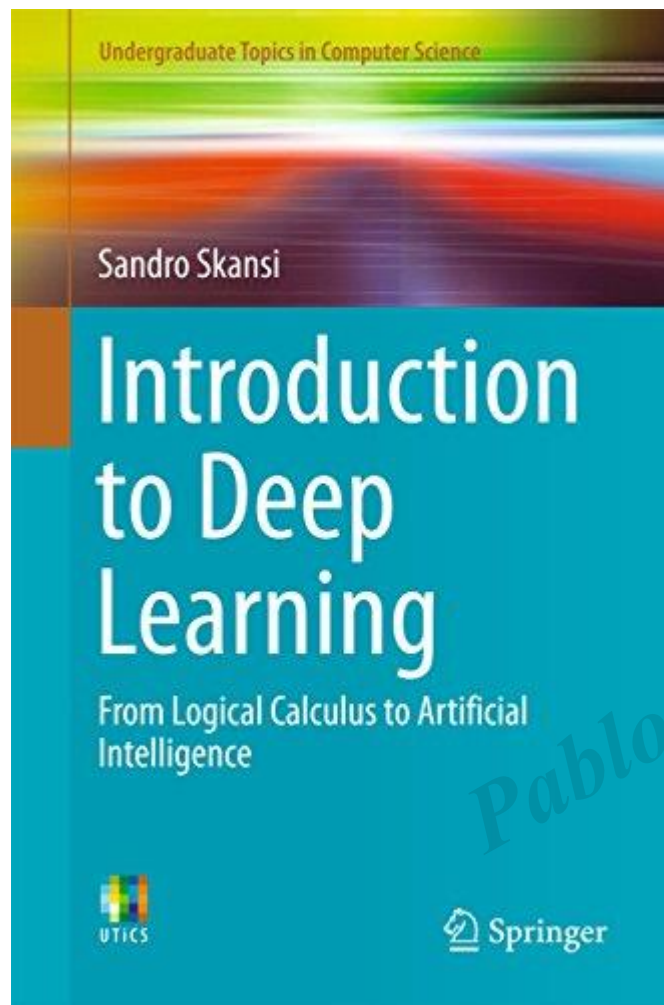
DEEP LEARNING

Prof. Dr. Jeronimo Marcondes

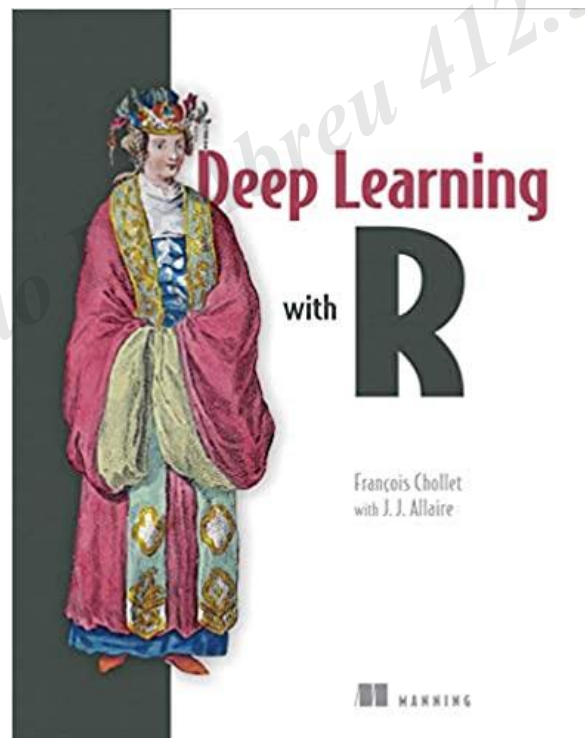
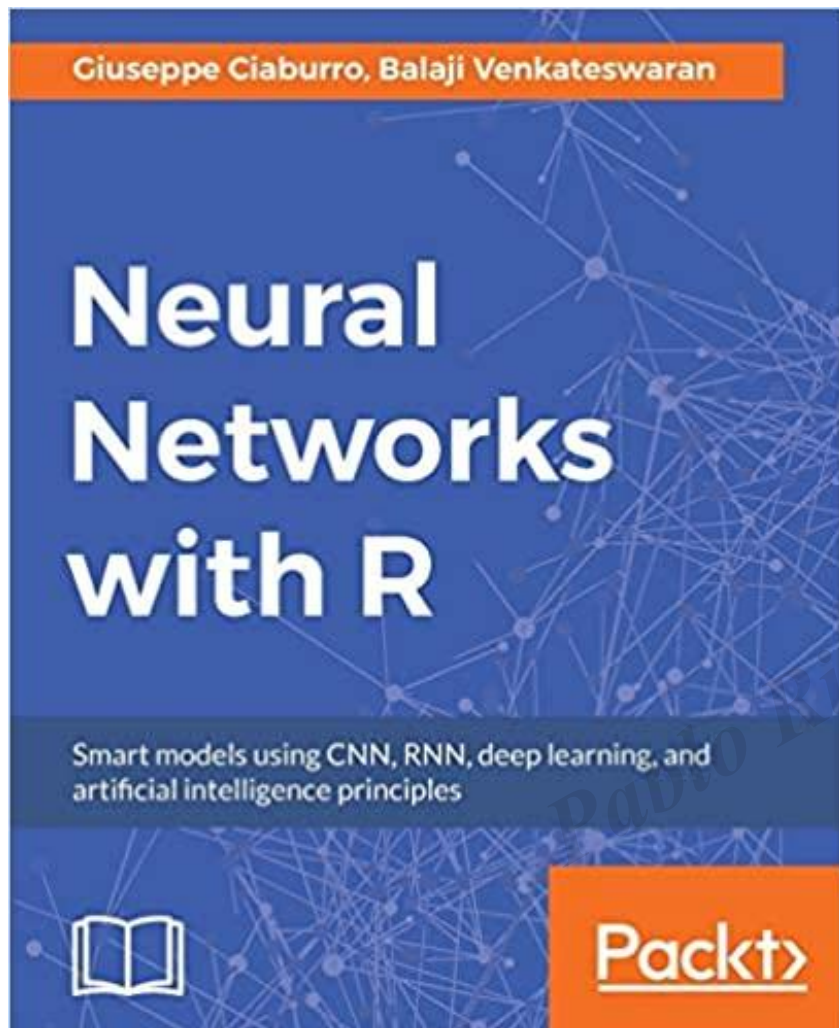
*El profesor es responsable de la idoneidad, originalidad y legalidad de los contenidos didácticos presentados.

Prohibida la reproducción total o parcial, sin autorización. Ley nº 9610/98

Introducción



Introducción



Introducción

- Deep Learning no supervisado
- ¿Qué es no supervisado?
- ¿Cuál es la diferencia con los modelos de FNN?

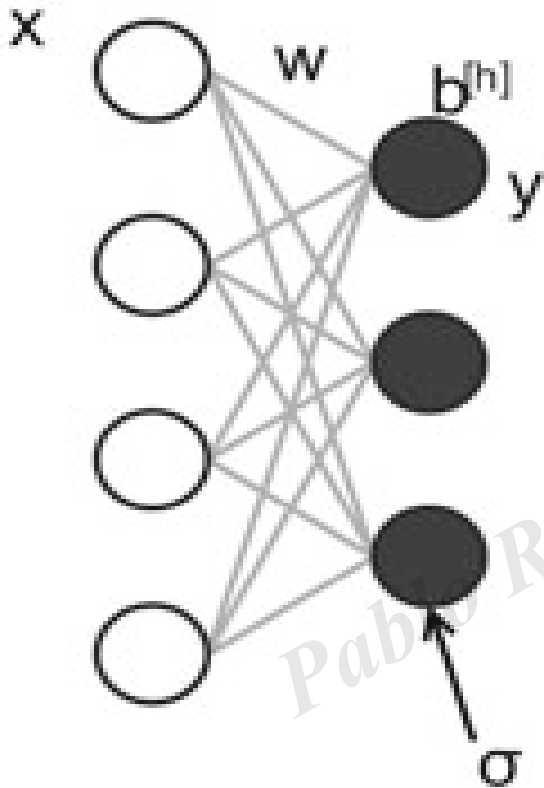
Introducción

- Modelos Basados en Energía
- Objetivo: Reducir la energía
- Similar al problema con nuestra función costo

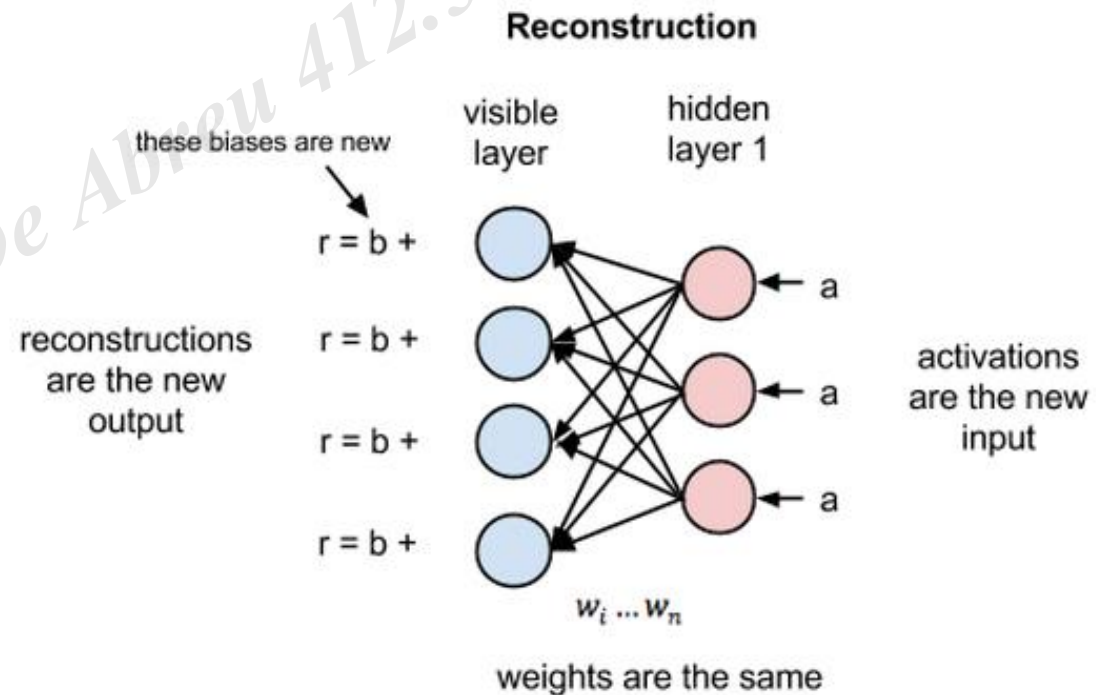
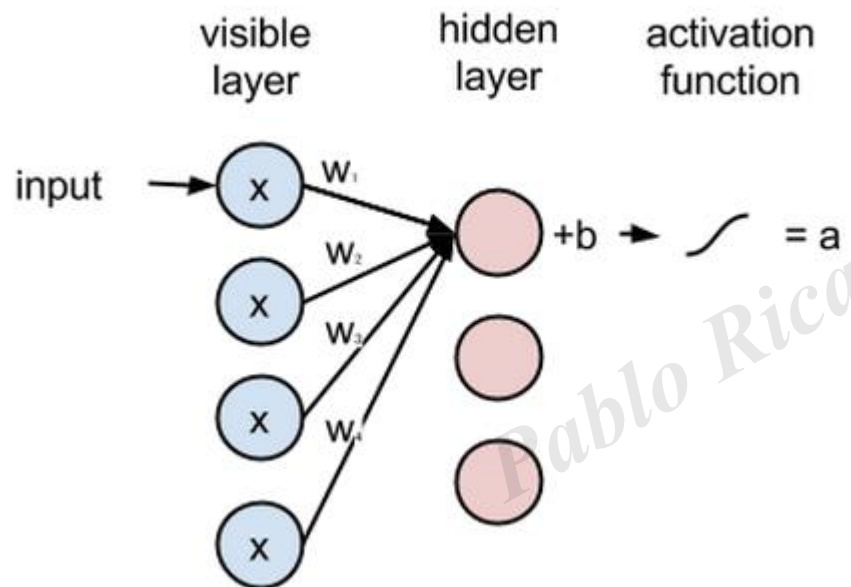
Máquinas de Boltzmann

- La máquina de Boltzmann es una FNN de una capa
- Nuestro objetivo: aprender la distribución de probabilidades de los inputs
- Ajustar pesos para que sea posible reconstruir los inputs (entradas)
- Máquina Restrita de Boltzmann

Máquinas de Boltzmann



Máquinas de Boltzmann



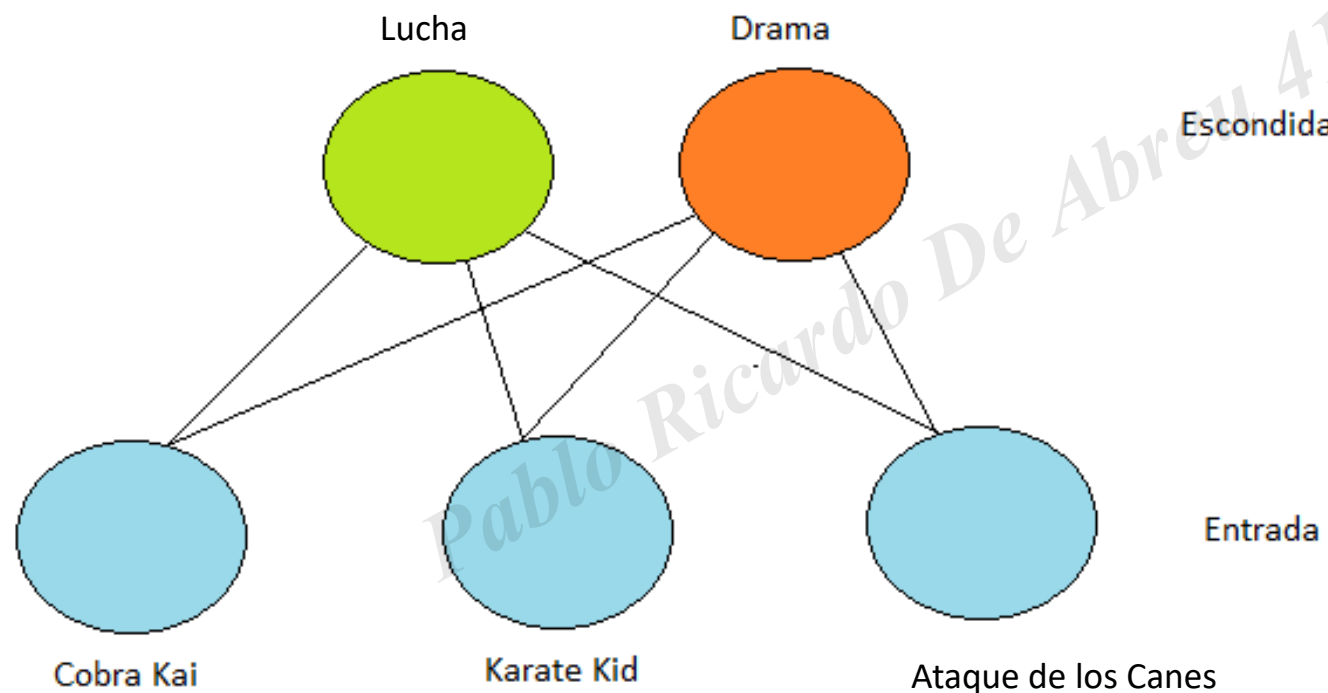
Fuente: <https://wiki.pathmind.com/restricted-boltzmann-machine>

Máquinas de Boltzmann

Características:

- No posee capa de salida
- Entradas son pasadas para la capa oculta
- ¿Por qué es máquina “restricta”?

Máquinas de Boltzmann



Máquinas de Boltzmann

Funcionamiento:

- Inicia capa interna aleatoriamente.
- Recibe insumo x .
- Calcula (probabilidad condicional que la neurona sea activada):

$$y = \sigma(xw + b_h)$$

Máquinas de Boltzmann

Funcionamiento:

- Y es devuelto para la capa interna para reconstrucción.

- Calcula:

$$r = \sigma(yw + b_v)$$

- Intuición: variable latente

Máquinas de Boltzmann

Aprendizaje por la Divergencia Contrastiva

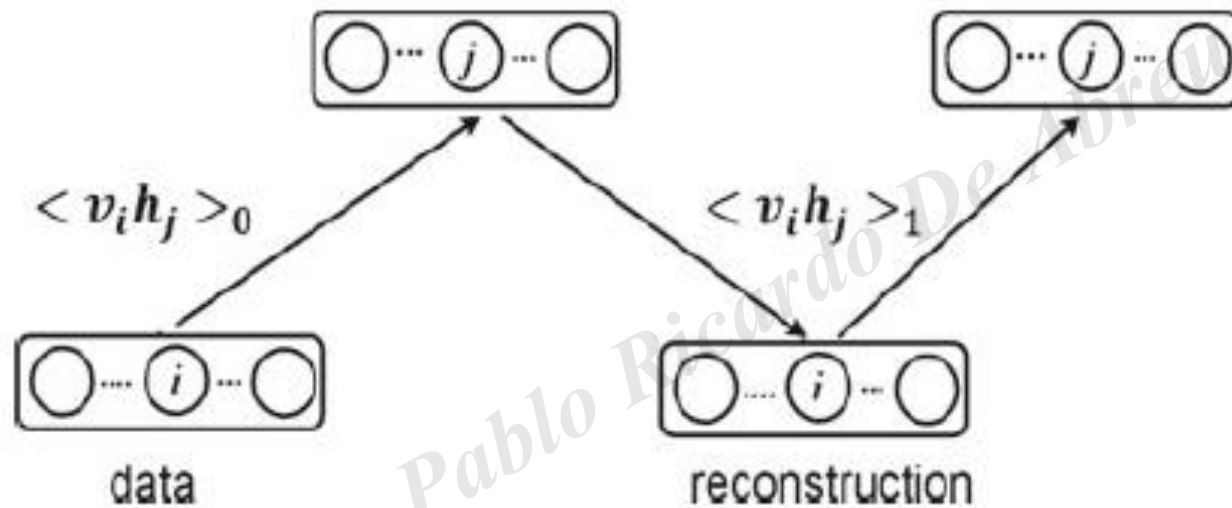
La divergencia contrastiva es una técnica de entrenamiento alternativa para aproximar la inclinación gráfica que representa la relación entre los pesos de una red y su error, llamada de gradiente. Como la mayoría de los algoritmos de aprendizaje probabilístico intenta optimizar el valor de la probabilidad logarítmica, ese gradiente representa la dirección deseada de cambio, de aprendizaje, para los parámetros de la red.

Máquinas de Boltzmann

Funcionamiento:

- La diferencia entre entrada y r es verificada por la divergencia de Kullback-Leibler.
- Fase negativa y positiva

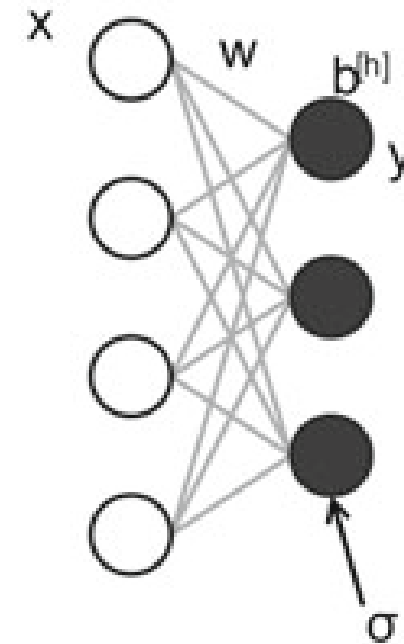
Máquinas de Boltzmann



Máquinas de Boltzmann

Aplicaciones:

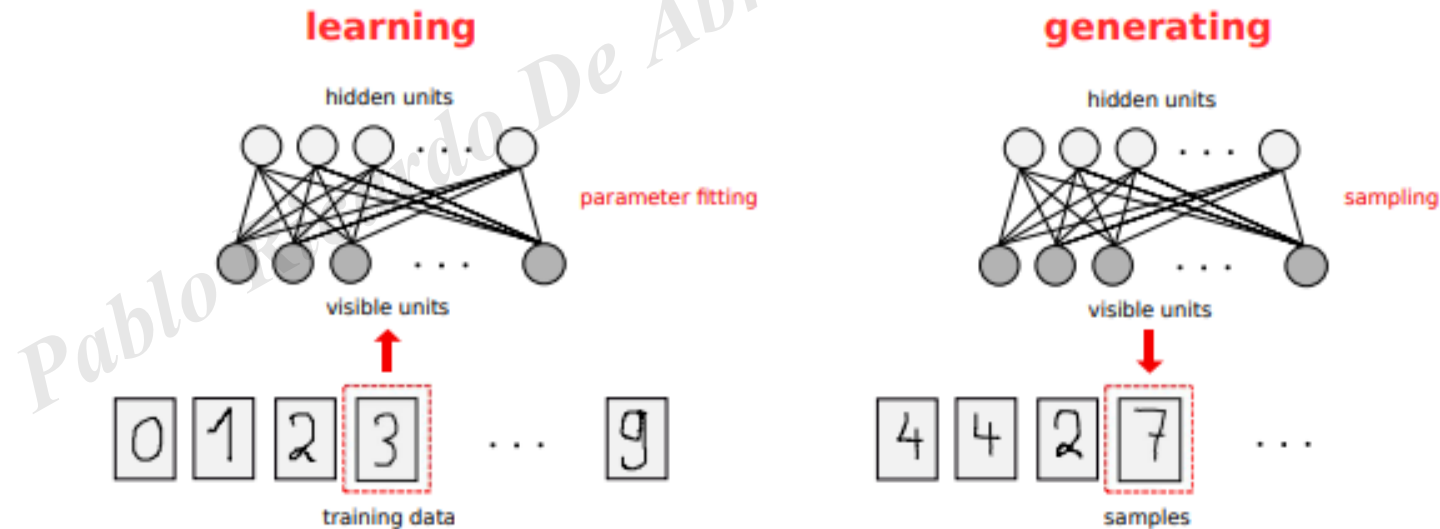
- Reducción de Dimensionalidad
- Ejemplo: problema con muchas variables.



Máquinas de Boltzmann

Aplicaciones:

- Sistemas de recomendación
- Reconstrucción de imagen



Máquinas de Boltzmann

Ejemplo: caso de NETFLIX

- Recomendación basada en contenido – elija similar a lo que ya ha hecho
- Recomendación basada en filtrado colaborativo – perfiles semejantes

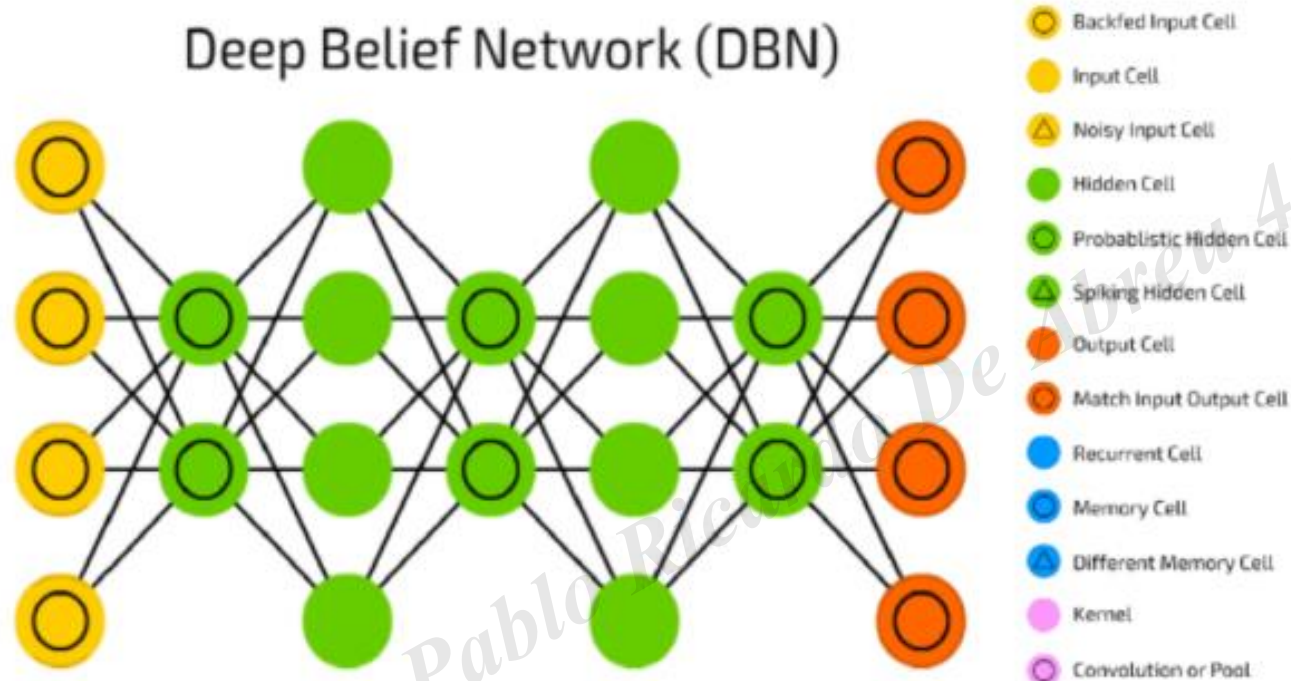
<https://tecnetit.com.br/como-a-netflix-usou-a-data-science-para-melhorar-seu-sistema-de-recomendacao/#:~:text=Para%20sua%20solu%C3%A7%C3%A3o%20eles%20usaram,dimensional%20ideal%20para%20seus%20usu%C3%A1rios.>

DBN

Deep Belief Networks

- Caso más generalizado de la Máquina restricta de Boltzmann.
- Apilar máquinas restrictas de Boltzmann

DBN

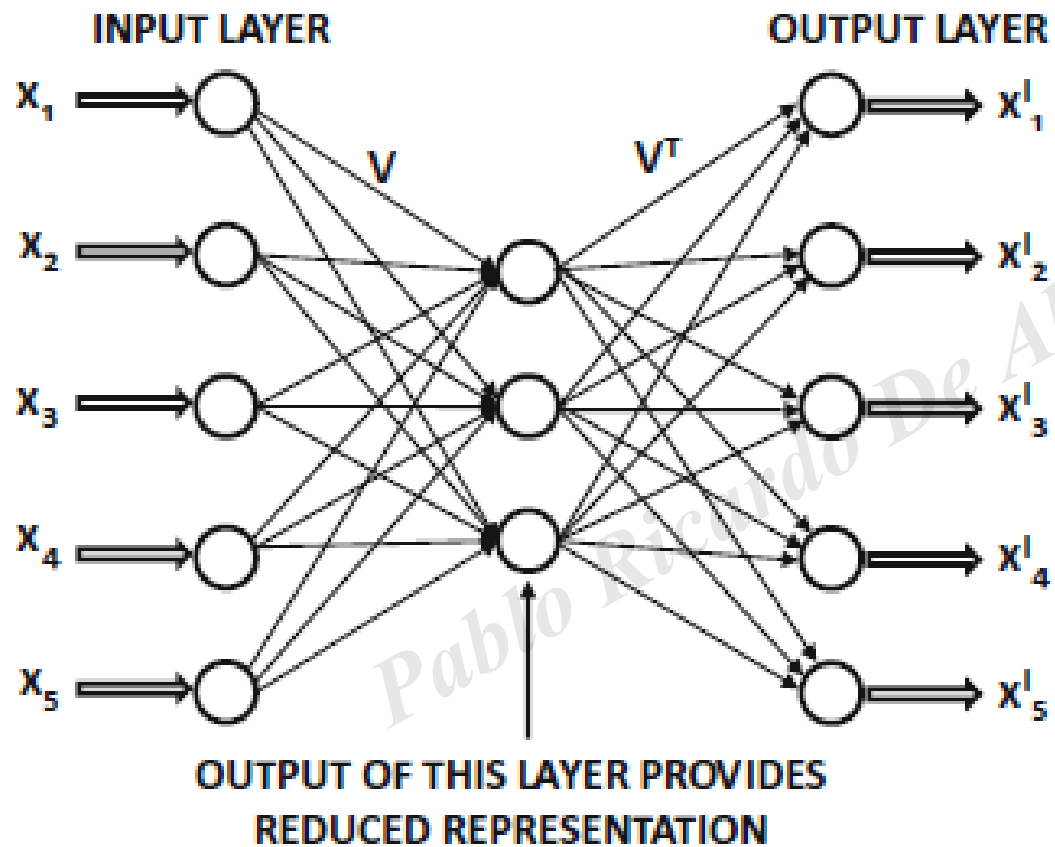


Fuente: <https://www.deeplearningbook.com.br/>

Autoencoder

- Objetivos muy semejantes a la Máquina de Boltzmann
- Reconstrucción de input
- Reducción de dimensionalidad

Autoencoder



Autoencoder

- Autoencoder es como un embudo
- Pasa la información y la restringe en un número menor de capas
- Después de eso intentamos reconstruir la información

Autoencoder

- Etapa de “encode” – codifica la imagen en una dimensionalidad menor:

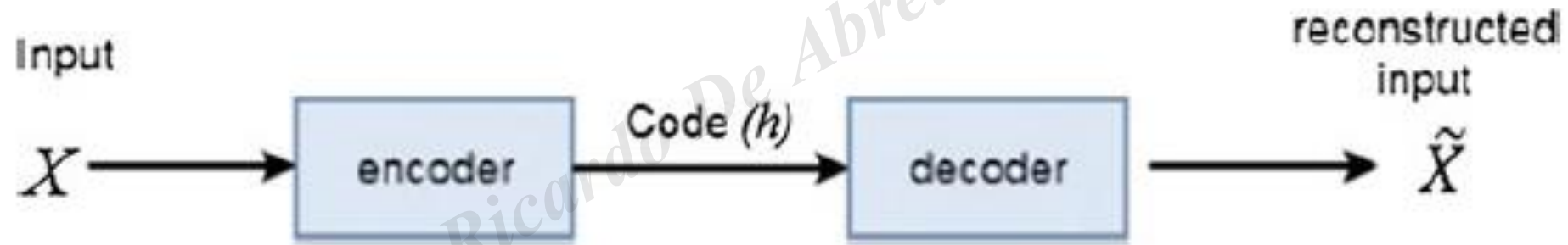
$$h = f(x)$$

- Etapa de “decode” – decodifica la imagen de manera que reconstruya el input.

$$r = f(h)$$

- Diferencia con relación a la máquina restringida de Boltzmann

Autoencoder

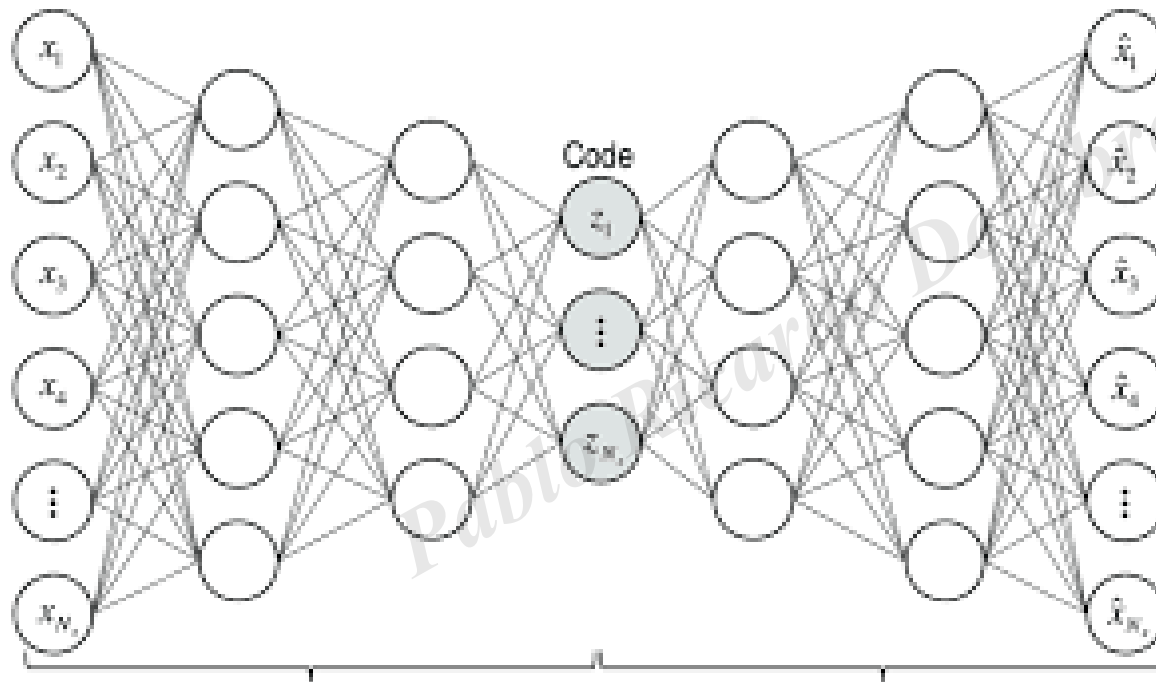


Autoencoder

<https://douglasduhaime.com/posts/visualizing-latent-spaces.html>

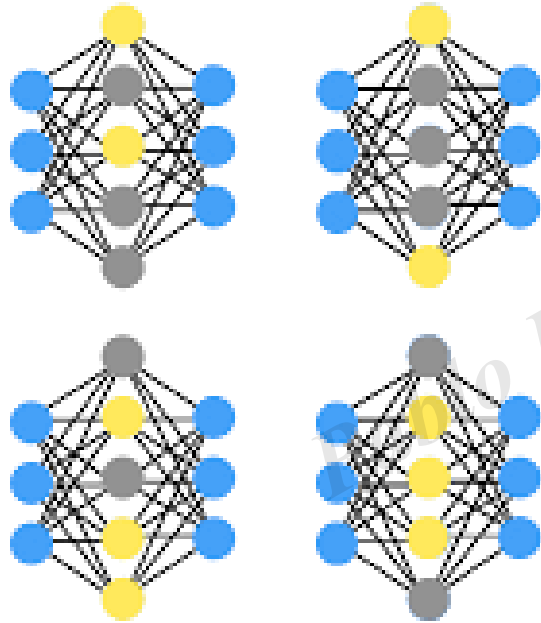
Tipos de Autoencoder

- Deep Autoencoder



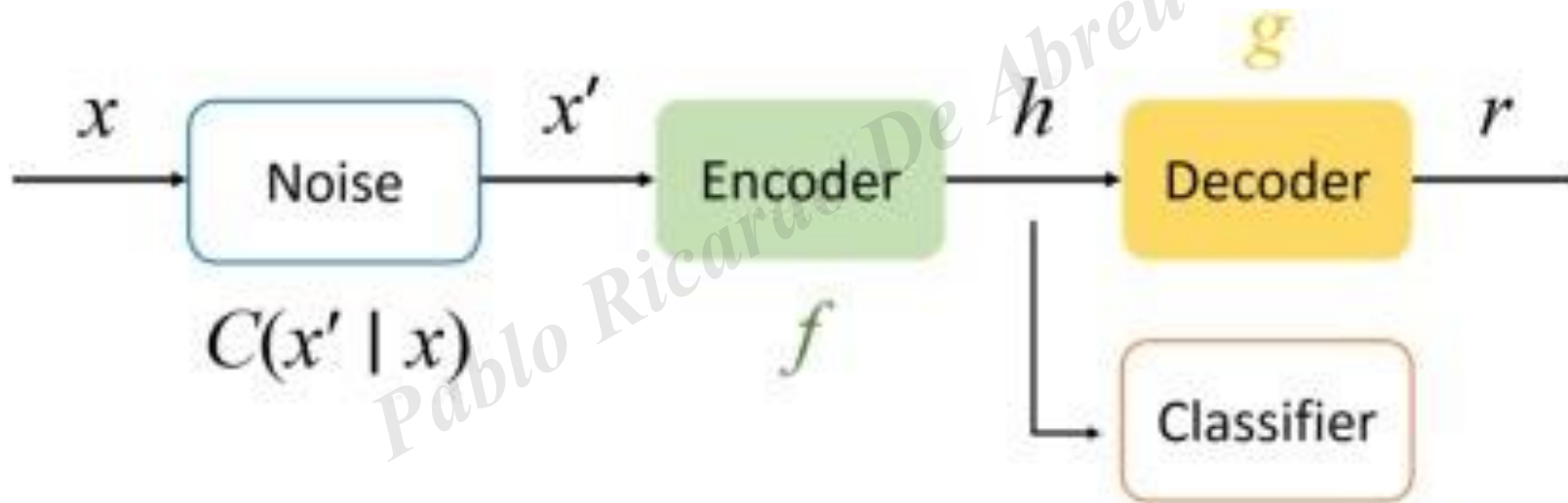
Tipos de Autoencoder

- Autoencoder Disperso
- Se aplica término de regularización



Tipos de Autoencoder

- Autoencoder Denoising



Tipos de Autoencoder

- Autoencoder Contractivo
- Diferencia con relación al denoising – incluye penalidad en la función costo
- Denoising incluye en los datos de entrenamiento
- Obtienen resultados más interesantes que el denoising

GAN

- Redes Adversarias Generativas
- “la idea más interesante en los últimos 10 años en Machine Learning”
- Dos redes compitiendo
- Modelo Generativo

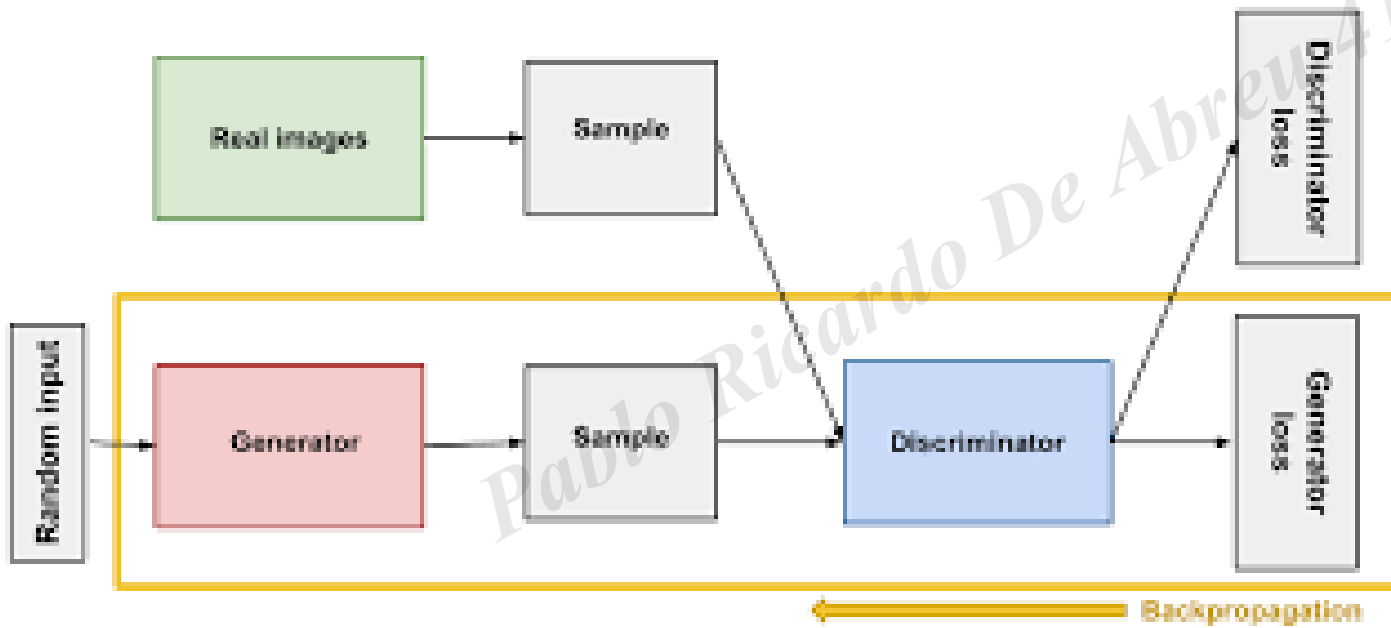
GAN



GAN

- ¿Cómo obtener y determinado x ?
- Modelos generativos modelan la distribución de clases individuales.
- Modelos discriminatorios aprenden la frontera entre clases.

GAN



GAN

- Ejemplo MNIST
- El generador va a generar imágenes “fake”
- Pasaremos imágenes reales junto con las generadas
- El discriminador debe reconocer las imágenes y hacer la diferencia

Parte Prática



GAN

- Realizamos el backpropagation
- El generador va a generar imágenes cada vez mejores
- El discriminador deberá quedar cada vez mejor para discriminar
- El discriminador obtiene imágenes reales y falsas y devuelve probabilidades, un número entre 0 y 1, con el 1 representando una previsión de imagen auténtica y con el 0 representando previsión de imágenes falsas (generadas por la red generativa).

Discussión

Pablo Ricardo De Abreu 412.566.518-42

¡GRACIAS!

<https://www.linkedin.com/in/jeronymo-marcondes-585a26186>