Machine Learning y aplicaciones en Física

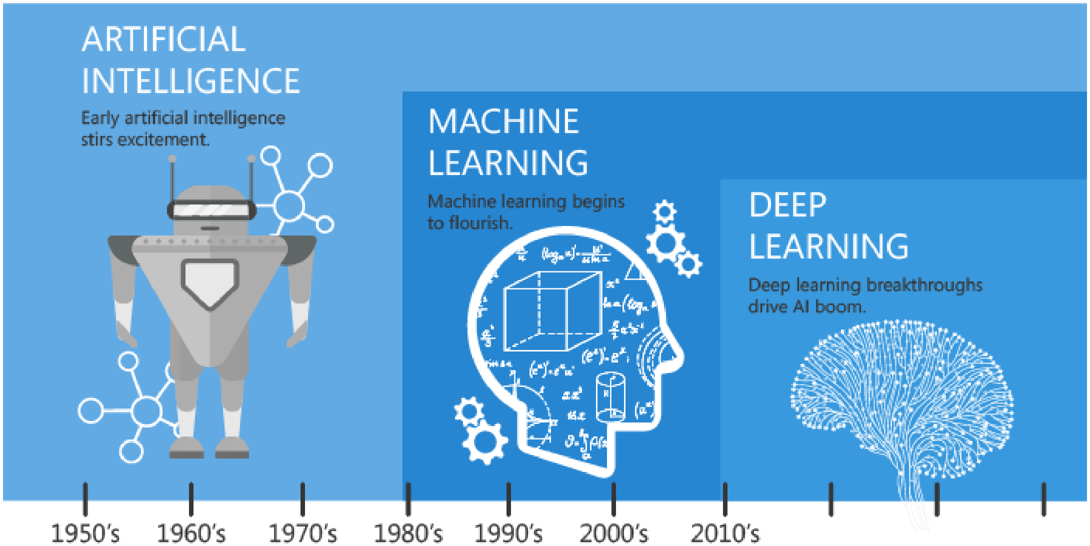
Diego A. Barbosa Trujillo

Universidad de los Andes

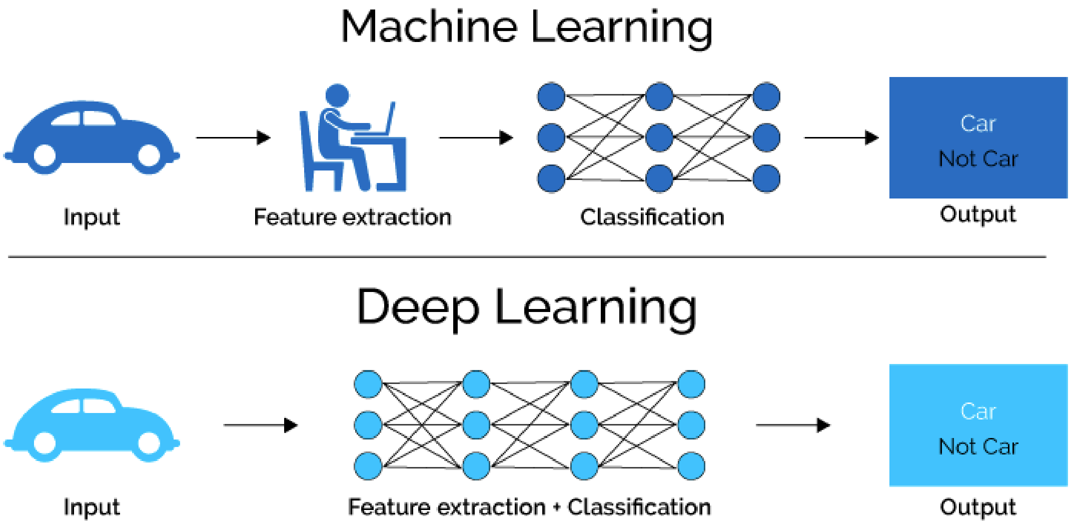
7 de febrero de 2020

Seminario de Mecánica Cuántica y Teoría de la Información.

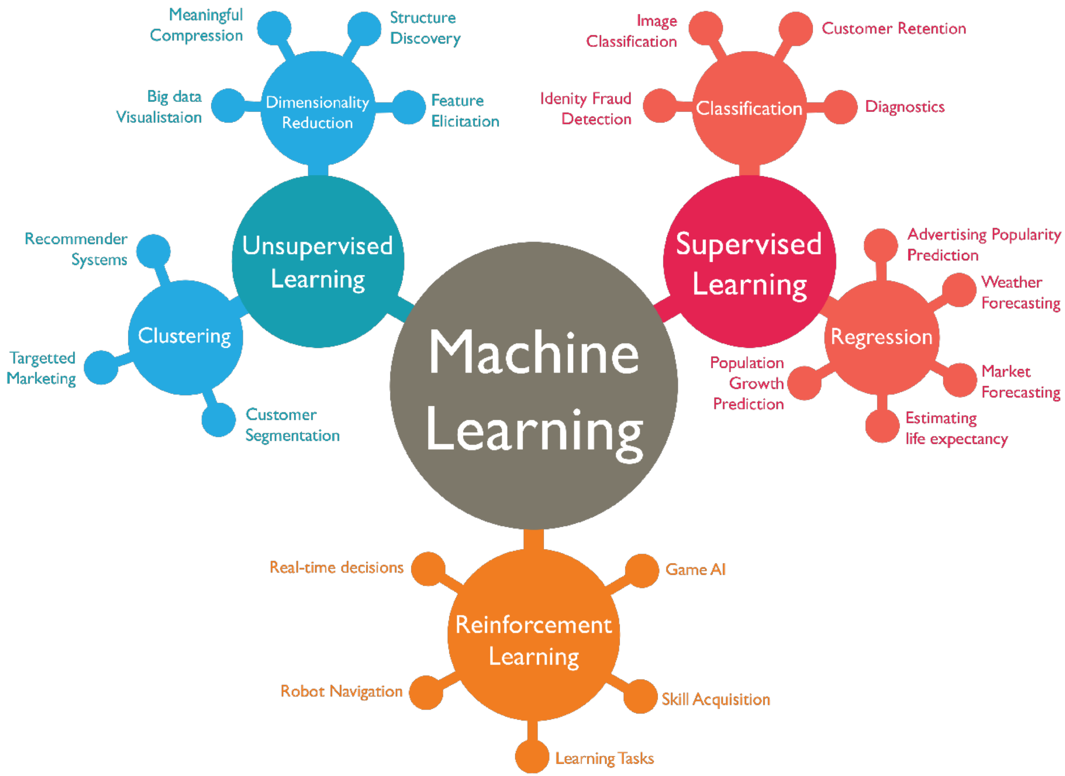
Machine Learning



- Arthur Samuel. IBM 1959.
- Hopfield, Rumelhart and Hinton. 1980.
- LeCun, Hinton and Bengio, 2010-Hoy.

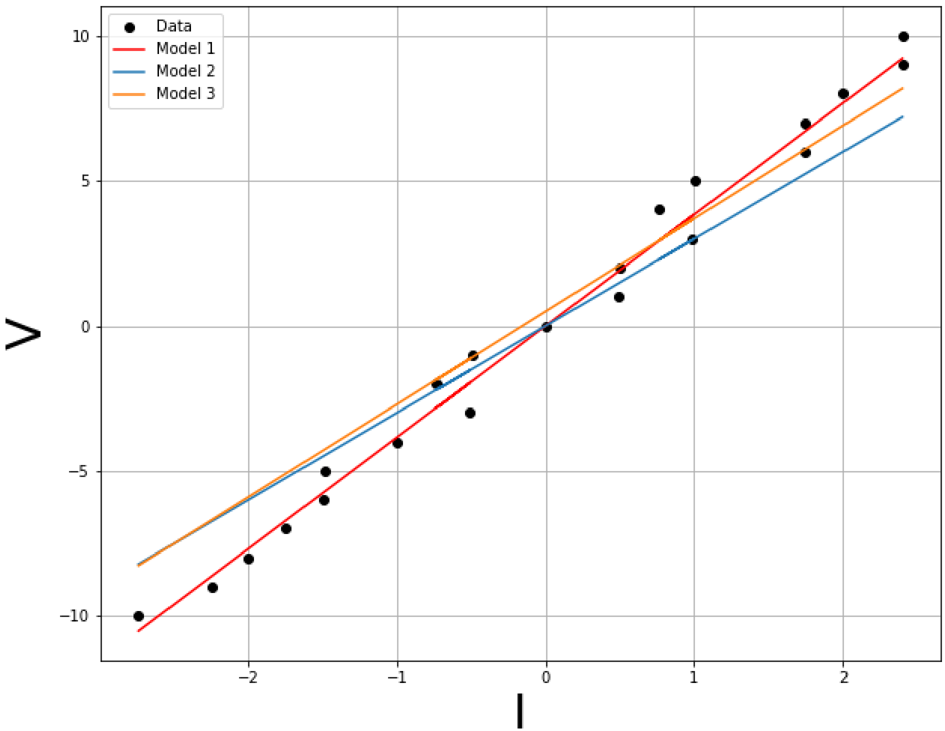

Machine Learning
Input
Feature extraction
Classification
Car
Not Car
Output
Deep Learning
Input
Feature extraction + Classification
Car
Not Car
Output

Machine Learning



Machine Learning

Machine Learning



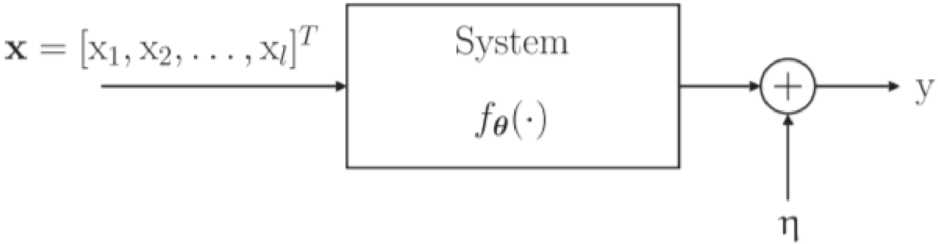
Regresión lineal

Regresión lineal

En los algorítmos de Machine Learning tendremos:

- Una serie de datos u observaciones: $\mathbf{X}$
- Un conjunto de funciones f_{θ^*}
- Una función de costo $\mathcal{C}$

Regresión lineal



La cual sabemos se puede expresar como:

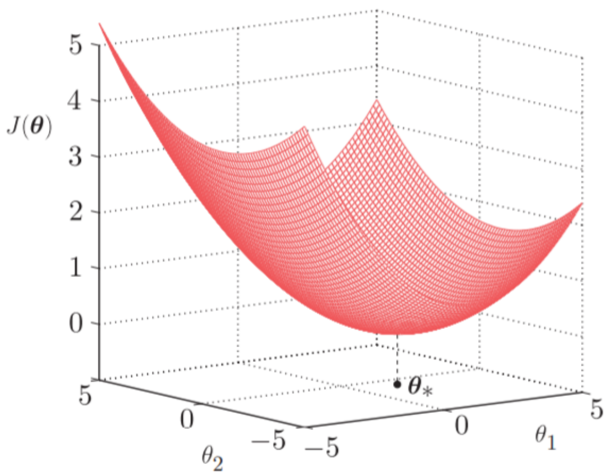
$$y = \boldsymbol{\Theta}^T \mathbf{x} + \eta$$

Regresión lineal

Para aprender los parámetros que describen las observaciones en regresión lineal, tenemos la función de costo de mímos cuadrados:

$$OLS = argmin_{\Theta}||\mathbf{y} - \mathbf{\Theta^T x}||_2^2 \tag{1}$$

la cual es una función convexa



Regresión lineal

Usando las reglas de cálculo y un poco de álgebra, encontramos una solución cerrada a este problema

$$\mathbf{\Theta}^* = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{2}$$

Sin embargo en el mundo real, la cantidad de datos es muy grande y para encontrar su solución se utilizan algorítmos como:

- Descenso de gradiente.
- Descenso de gradiente estocástico.

Regresión lineal

En muchas ocasiones la simple función de costo no se ajusta correctamente y debemos realizar una regularización sobre los parámetros.

$$\mathbf{\Theta}^*_{Ridge} = argmin_{\Theta}(||\mathbf{y} - \mathbf{\Theta}^{\mathbf{T}}\mathbf{x}||_2^2 + \lambda||\theta||_2^2) \tag{3}$$

o

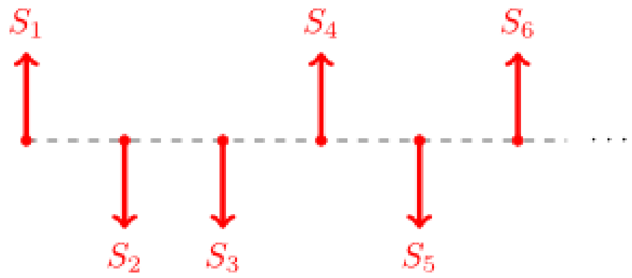
$$\mathbf{\Theta}^*_{Lasso} = argmin_{\Theta}(||\mathbf{y} - \mathbf{\Theta}^{\mathbf{T}}\mathbf{x}||_2^2 + \lambda||\theta||_1) \tag{4}$$

Aplicación en Mecánica Estadística

Regresión Lineal y Cadena de Ising 1D

Imaginemos que tenemos un ensemble de diferentes configuraciones de espines y su correspondiente energía en una dimensión, dada por:

$$H = -J \sum_{j=1}^{L} S_j S_{j+1} \tag{5}$$



Regresión Lineal y Cadena de Ising 1D

Supongamos que nos tenemos un conjunto de datos $D = (\{S_j\}_{j=1}^{L}, E_j)$ y nuestro objetivo es utilizar regresión lineal para encontrar la constante de acoplamiento J. Para este propósito propongamos un modelo de la forma

$$H_{modelo}[S^i] = \sum_{j}^{L} \sum_{k}^{L} J_{i,k} S_j^i S_k^i \tag{6}$$

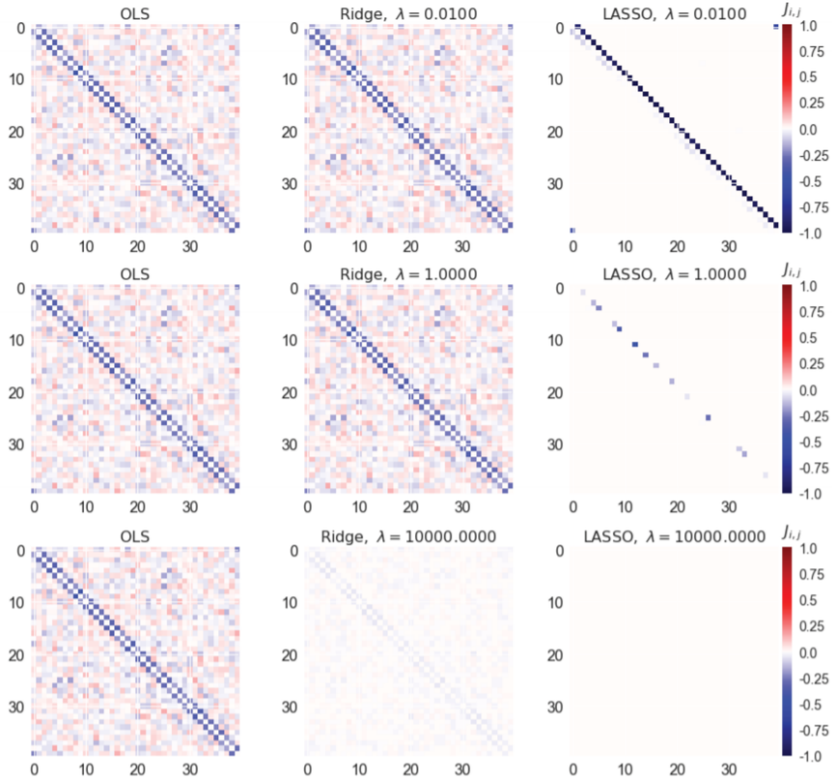
El cual podemos escribir como

$$H_{modelo}[S^i] = \mathbf{X}^i . \mathbf{J}, \tag{7}$$

donde X^i representa todas las interacciones de dos cuerpos $\{S_j^i S_k^i\}$ y el índice i corre sobre todos el conjunto de datos.

Regresión Lineal y Cadena de Ising 1D

Para este problema utilizamos 10000 cadenas aleatorias de Ising de longitud 40, y aplicando regresión lineal encontramos:



Clasificación

Regresión Logísitca

Regresión logística

Para entender la regresión logística, consideremos un sistema de dos estados cuyas energías son E_0 y E_1, la probabilidad de que el sistema esté en un estado de energía E_0 es

Regresión logística

Para entender la regresión logística, consideremos un sistema de dos estados cuyas energías son E_0 y E_1, la probabilidad de que el sistema esté en un estado de energía E_0 es

$$P(E_0) = \frac{\exp(-\beta E_0)}{\exp(-\beta E_0) + \exp(-\beta E_1)} = \frac{1}{1 + \exp(-\beta \Delta E)} \tag{8}$$

y para un estado de energía E_1

Regresión logística

Para entender la regresión logística, consideremos un sistema de dos estados cuyas energías son E_0 y E_1, la probabilidad de que el sistema esté en un estado de energía E_0 es

$$P(E_0) = \frac{\exp(-\beta E_0)}{\exp(-\beta E_0) + \exp(-\beta E_1)} = \frac{1}{1 + \exp(-\beta \Delta E)} \tag{8}$$

y para un estado de energía E_1

$$P(E_1) = 1 - P(E_0) \tag{9}$$

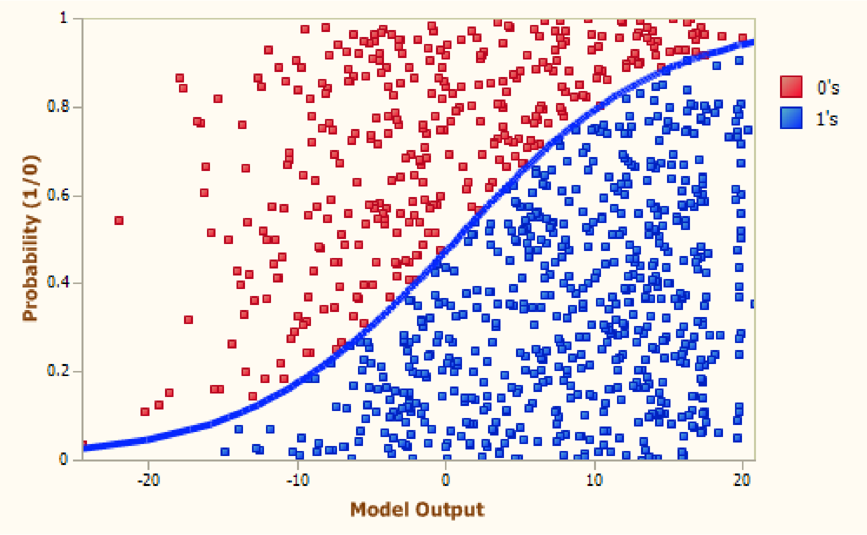
Regresión logística

Con el modelo de regresión logísitca lo que esperamos es realizar una clasificación de un conjunto de datos binario, a través de

$$P(y_i = 1|\mathbf{x}_i, \theta) = \frac{1}{1 + \exp(-\mathbf{x}_i^T \theta)} = \sigma(\mathbf{x}_i^T \theta)$$

y

$$P(y_i = 0|\mathbf{x}_i, \theta) = 1 - P(y_i = 1|\mathbf{x}_i \theta)$$



Regresión logística

Para realizar el entrenamiento de este clasificador, usaremos la función de costo

$$c(\theta) = -\sum_{i=1}^{n}(y_i \log \sigma(\mathbf{x}_i^T\theta) + (1-y_i)\log(1-\sigma(\mathbf{x}_i^T\theta)).$$

También conocida como entropía cruzada.

Regresión Logísitca y Modelo de Ising

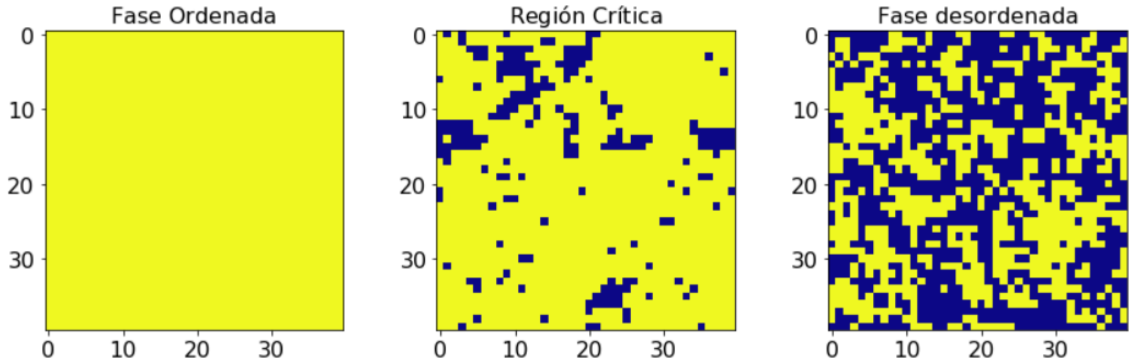
Clasificación

Usando el modelo de regresión logística, podemos realizar una clasificación de los estados de Ising de acuerdo a su fase. Recordemos que el Hamiltoniano del modelo de Ising clásico es

$$H = -J \sum_{<ij>} S_i S_j,$$

con $S_j \in \{\pm 1\}$ y J la constante de interacción.

Clasificación



- Tenemos una fase ordenada (ferromagnética).
- Onsager demostró que que este modelo sufre un transición de fase a una temperatura crde T_c/J approx 2.26.
- Fase desordenada.

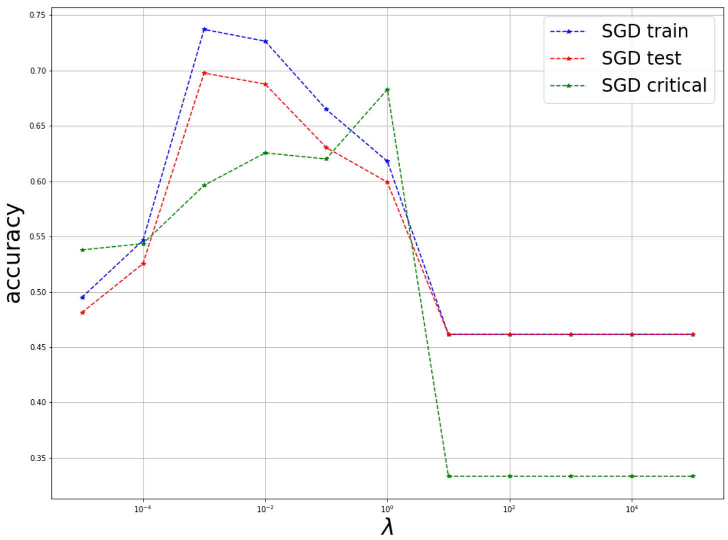
Clasificación

Para aplicar regresión logística al modelo de Ising para las dos fases:

- Consideramos un lattice cuadrado de tamaño 40x40 y usando montecarlo se preparan 10^4 estados variando la temperatura.
- Para entrenar el clasificador, usamos descenso de gradiente estocástico usando configuraciones con temperaturas $T/J < 2{,}0$, $2{,}0 \leq T/J \leq 2{,}5)$ y $T/J > 2{,}4$.
- Una vez entrenado el clasificador, evaluamos el desempeño del clasificador en estados de prueba.

Clasificación

La siguiente gráfica muestra los resultados, variando un hiperparámetro del gradiente de descenso estocástico



Conclusiones

- Machine Learning es una herramienta poderosa incluso con modelos simples. Los cuales pueden ser utilizados en física para extraer información.
- En muchos de los modelos como regresión logística y redes neuronales existe una conexión con la física estadística.