

UNIVERSIDAD  
PANAMERICANA

# Machine Learning

(<https://leonpalafox.github.io/mlclase/>)

Leon F. Palafox PhD

# Noticias

MIT  
Technology  
Review

[Log In / Create an account](#) [Search q](#)

[Topics+](#) [The Download](#) [Magazine](#) [Events](#) [More+](#)

[Subscribe](#)



MATTHEW BARNES / GETTY

---

Intelligent Machines

---

## Machine learning predicts World Cup winner

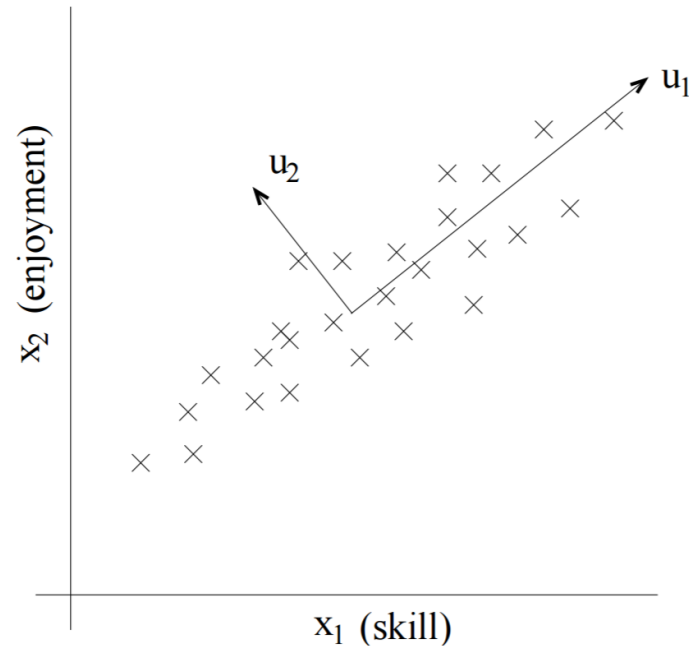
Leon Palafox

# Escenario

- Estamos tratando de detectar si alguien es buen conductor o mal conductor.
- Hacemos una encuesta:
  - Medimos distintas variables:
    - Skill (basado en metricas)
    - Diversion (basado en tiempo)
    - Precisión
    - Edad, genero, etc

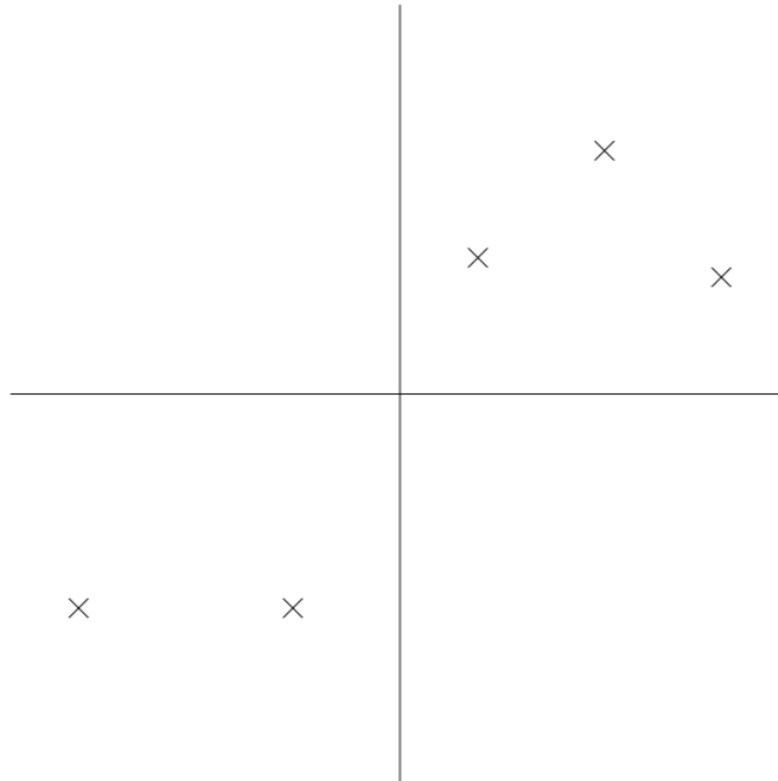
# Hay variables que sobran

- Creamos una nueva variable
  - Esta variable es el “karma” de cada usuario
  - La nueva variable captura la varianza

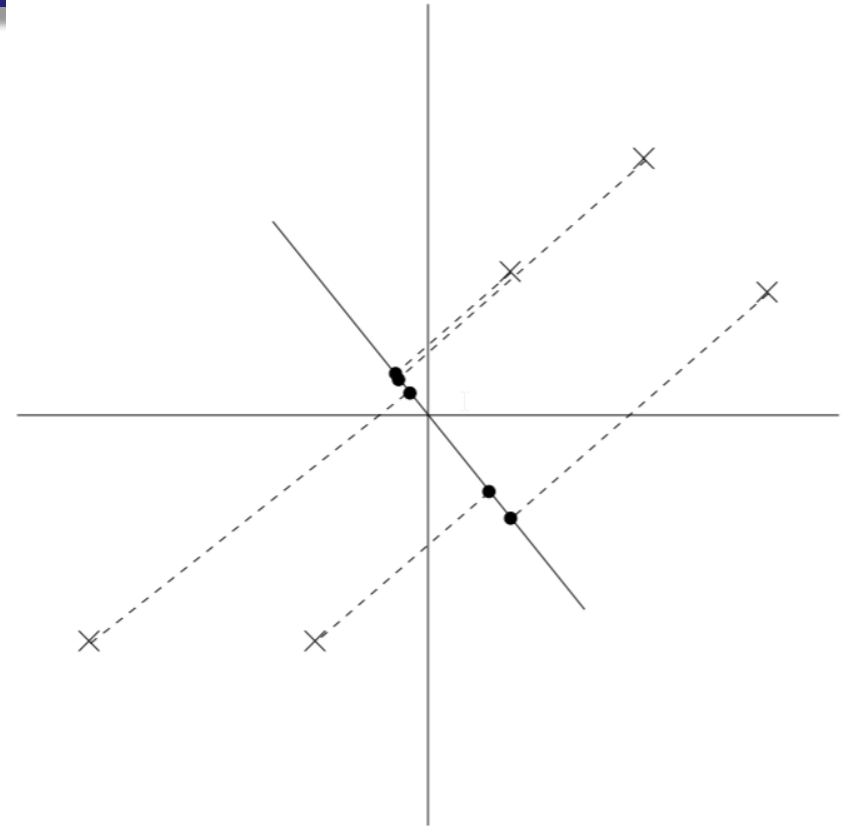
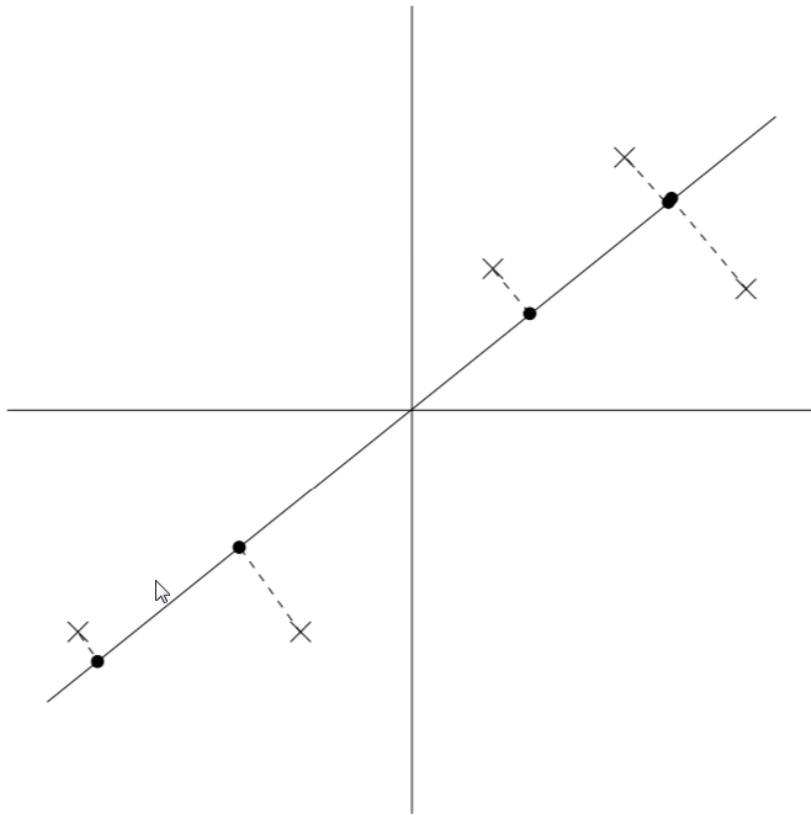


# PCA

- Buscar la dirección donde los datos varían mas



# Options, options!



# Descomposición de Features

- Cuando tenemos muchos features
  - Muchos son redundantes
    - Velocidad en diferentes dimensiones
    - Electrodo en gorros de EEG
    - Edad y Fecha de Nacimiento
- Es muy común utilizar la matriz de correlaciones
  - Esto nos indica que Features están correlacionados y los podemos eliminar.

# ICA (Análisis de Componentes Principales)





# Definición del problema

- Dadas  $s$  (fuentes), una matriz de transformación  $A$ , la señal observada es  $x$

$$x = As,$$

- El problema consiste en encontrar  $s$  y  $A$ .

# Algunas limitaciones

- Es imposible encontrar las escalas correctas.
- Los datos no pueden tener una distribución normal.

# Computacionalmente

- ICA es mucho más intensivo que PCA, ya que trata de encontrar  $s$  y  $A$
- En general debemos utilizar Normalización

# Normalización

1. Let  $\mu = \frac{1}{m} \sum_{i=1}^m x^{(i)}$ .
2. Replace each  $x^{(i)}$  with  $x^{(i)} - \mu$ .
3. Let  $\sigma_j^2 = \frac{1}{m} \sum_i (x_j^{(i)})^2$
4. Replace each  $x_j^{(i)}$  with  $x_j^{(i)} / \sigma_j$ .

# Aprendizaje de Manifold

- Conforme tenemos más dimensiones, visualizar los datos se vuelve más difícil
- Aprendizaje de manifold se utiliza para analizar datos
- Es importante para otras áreas de ML

# T-SNE

t-distributed stochastic neighbor embedding

<http://everynoise.com/engenremap.html>

<http://cs.stanford.edu/people/karpathy/tsnejs/csvdemo.html>

# T-SNE theory

- Dados dos espacios X y Y donde:
  - Dimensionalidad de X >> Dimensionalidad de Y (2-3)
- T-SNE calcula distribuciones que garanticen que objetos similares van a estar cerca.

$$p_{j|i} = \frac{\exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|\mathbf{x}_i - \mathbf{x}_k\|^2 / 2\sigma_i^2)},$$

$$q_{ij} = \frac{(1 + \|\mathbf{y}_i - \mathbf{y}_j\|^2)^{-1}}{\sum_{k \neq i} (1 + \|\mathbf{y}_k - \mathbf{y}_i\|^2)^{-1}}$$

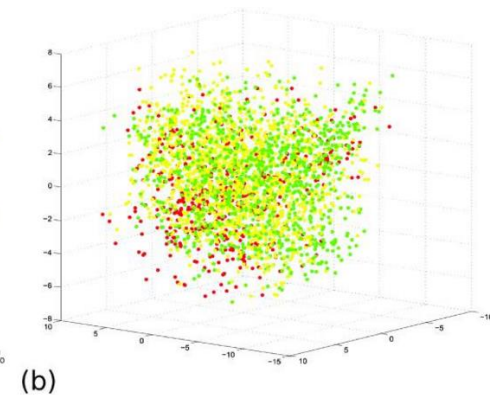
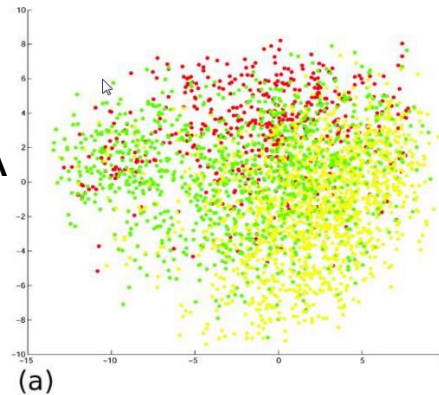
$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2N}$$

# Aplicaciones

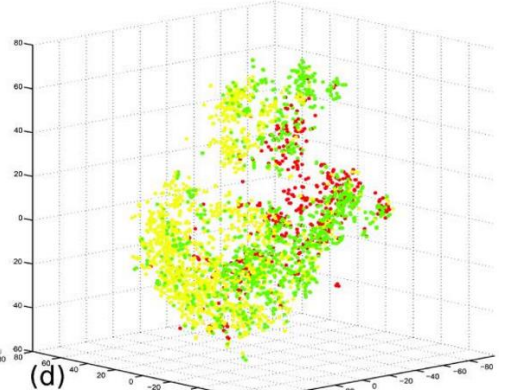
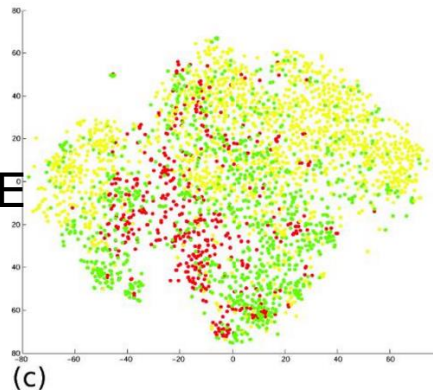
- Cancer Prognosis

● Malignant    ● Cystic    ● Benign Mass

PCA



T-SNE

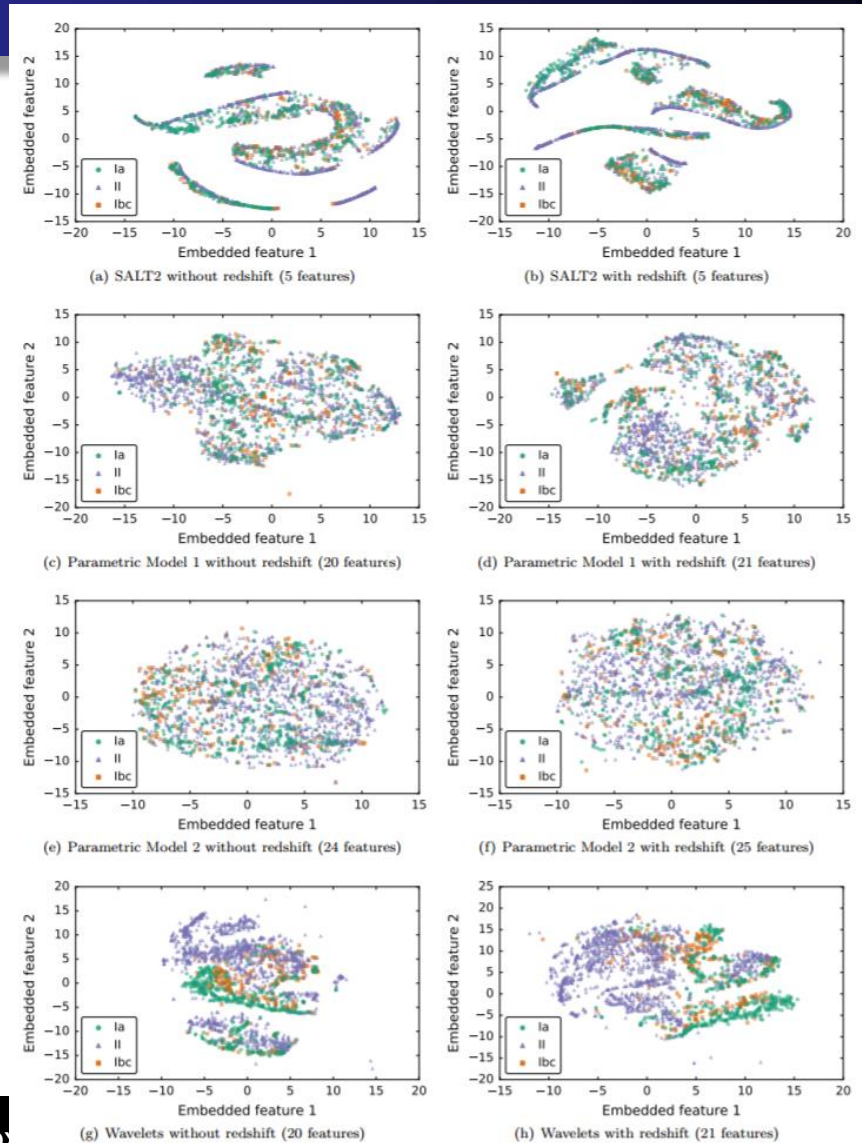


Jamieson, Andrew R., et al. "Exploring nonlinear feature space dimension reduction and data representation in breast CADx with Laplacian eigenmaps and t-SNE." *Medical physics* 37.1 (2010): 339-351.

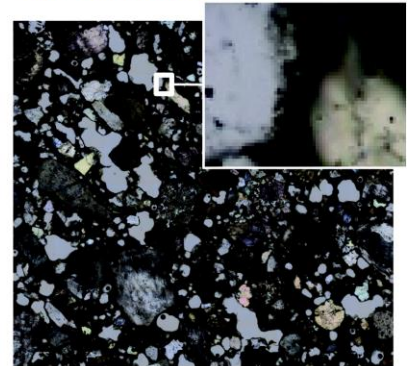
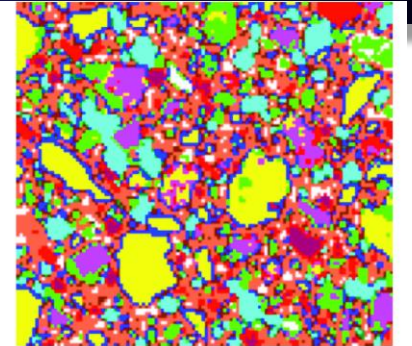
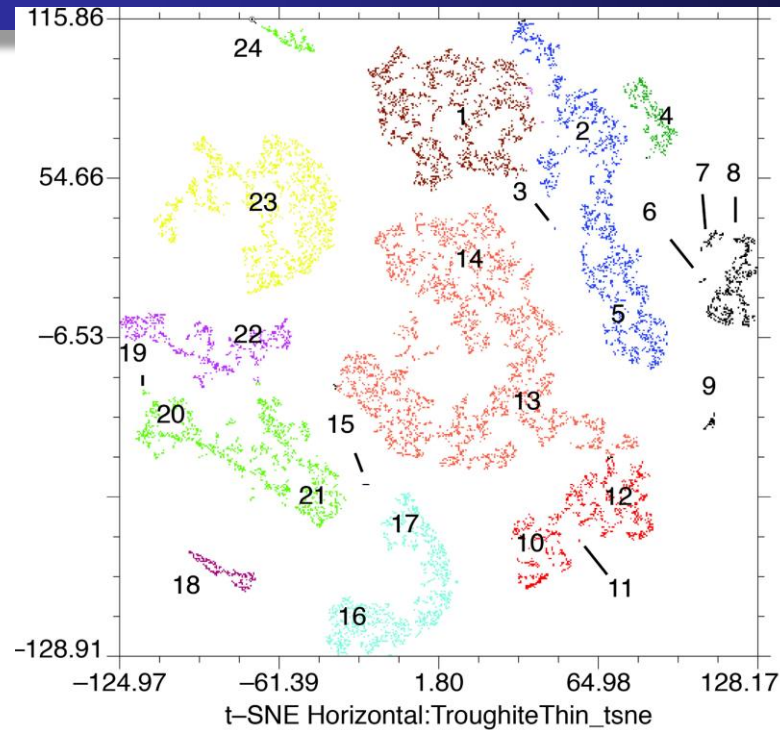


# Aplicaciones

Lochner, Michelle, et al. "Photometric Supernova Classification with Machine Learning." *arXiv preprint arXiv:1603.00882* (2016).



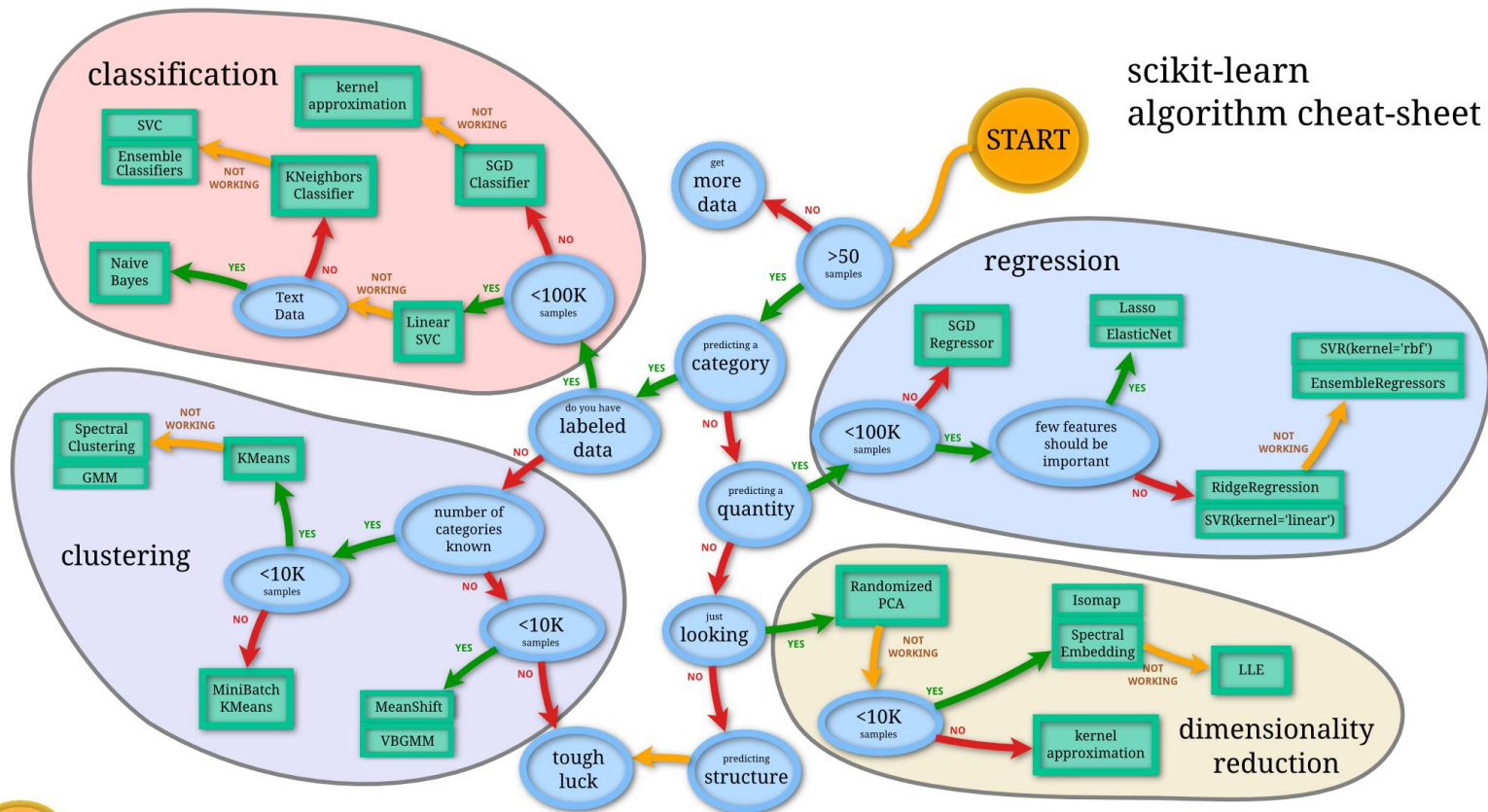
# Aplicaciones



Thompson, David R., et al. "Automating X-ray Fluorescence Analysis for Rapid Astrobiology Surveys." *Astrobiology* 15.11 (2015): 961-976.

# Cheat Sheet

scikit-learn  
algorithm cheat-sheet



[http://scikit-learn.org/stable/tutorial/machine\\_learning\\_map/](http://scikit-learn.org/stable/tutorial/machine_learning_map/)

Leon Palafox

# Futuro de ML

- ML tiene un future prometedor como industria
- Deep Learning esta comenzando a crear muchas industrias
- Apple esta comenzando a desarrollar muchas cosas con ML
- Muchas personas concuerdan en que ML es el principio de la nueva revolución industrial.

# Consecuencias

- Aún mas especialización
  - Mas desempleo
- Quienes no lo tomen en serio, pueden caer en la obsolescencia muy rapido.
  - Liverpool, Palacio de Hierro
- Expertos en ML van a ser muy demandados.

# Objetivo de la clase

- Familiarizar a las personas con ML
- Mostrar como diferentes sets de datos tienen diferentes problemas
- Mostrar que ML no es Plug and Lay y si require conocimiento acerca de los algoritmos
- Emocionarlos acerca de ML