

# Machine Learning for Geosciences

By Francisco Mendoza

[mentofran@gmail.com](mailto:mentofran@gmail.com)



Course material:

<https://github.com/mathphysmx/teaching-ml>

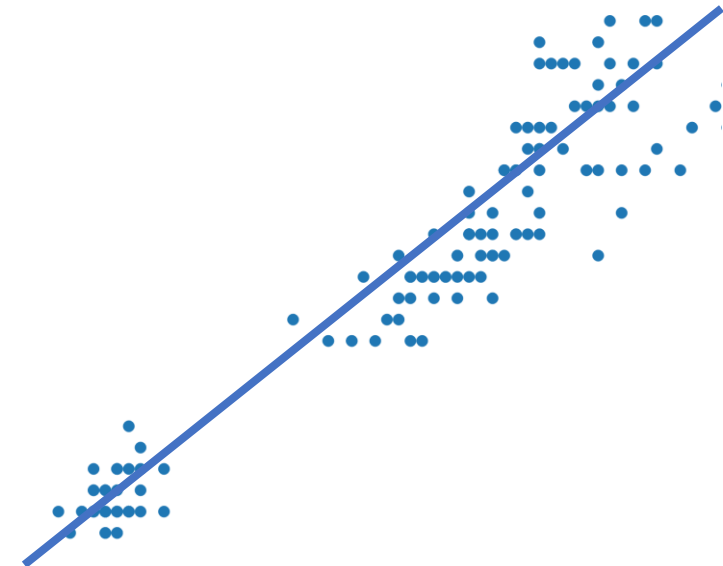
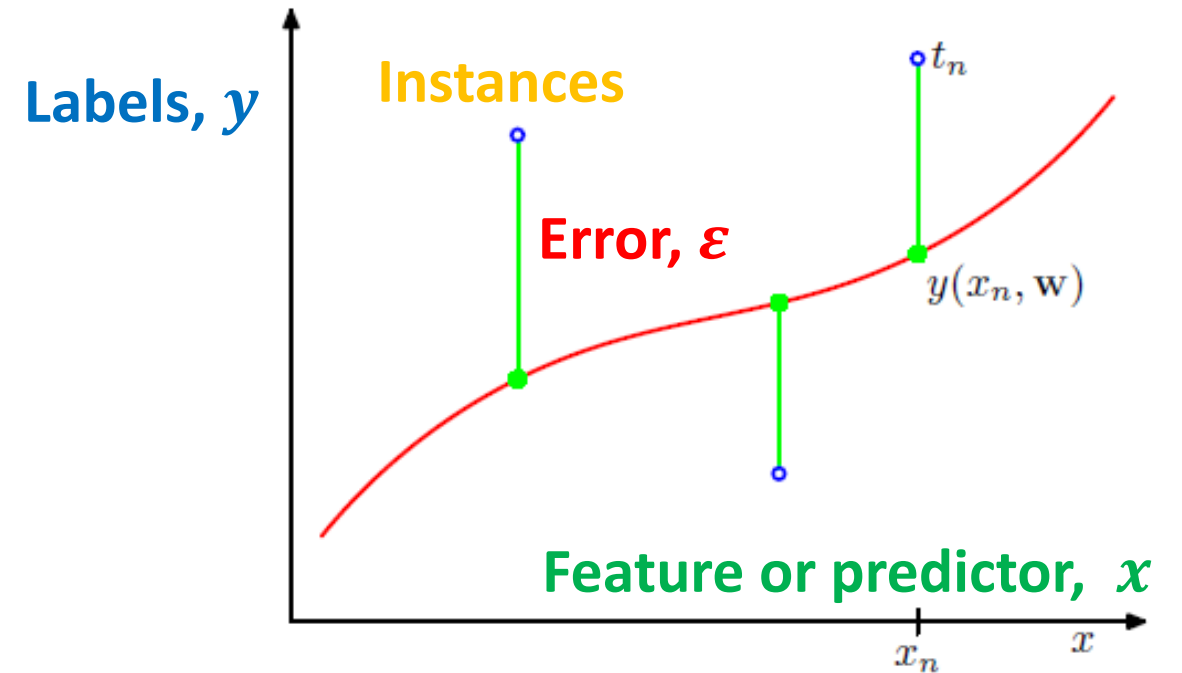
# Basic concepts

All previous concepts used in statistics such as (in)?dependent variable, input/output variables,...

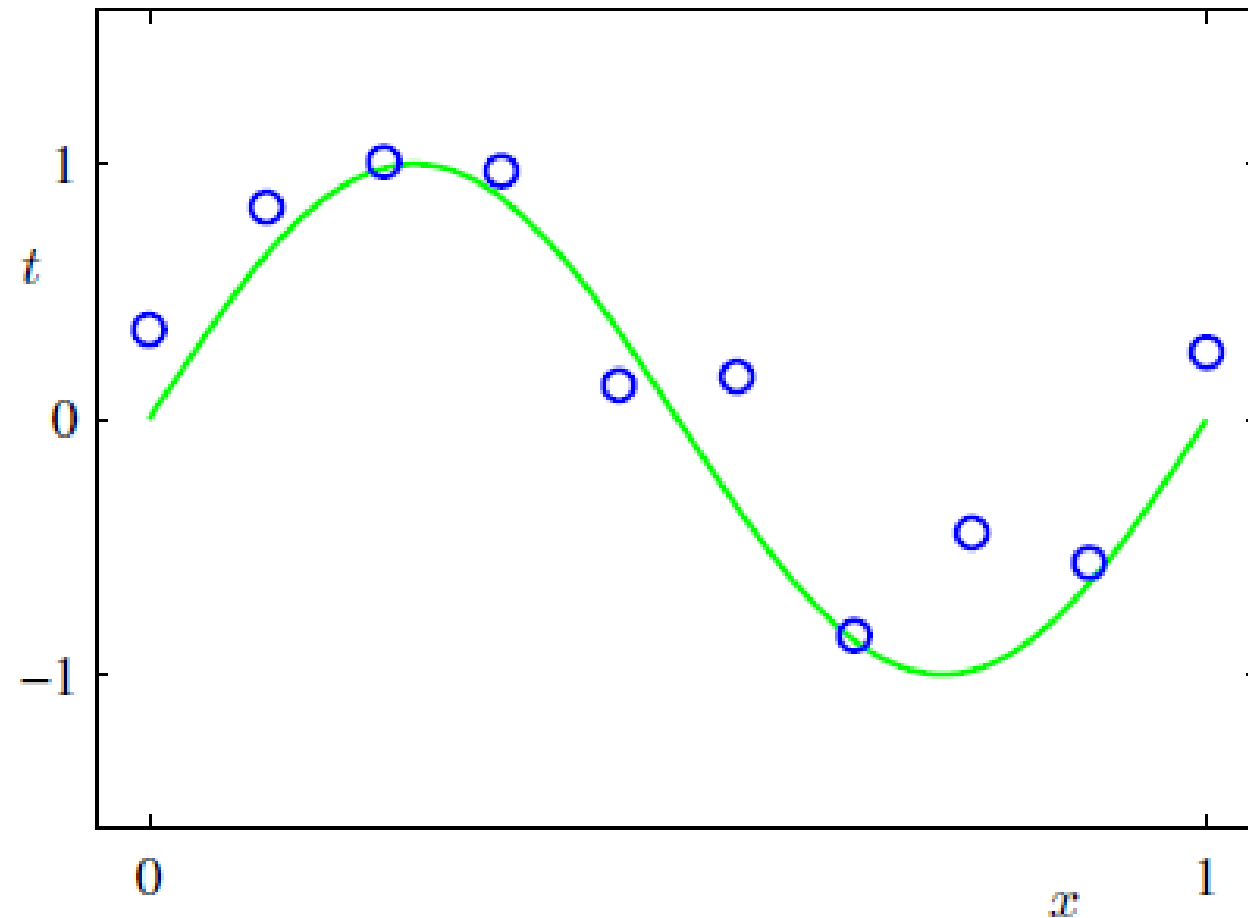
Training set

Features   Labels   Error

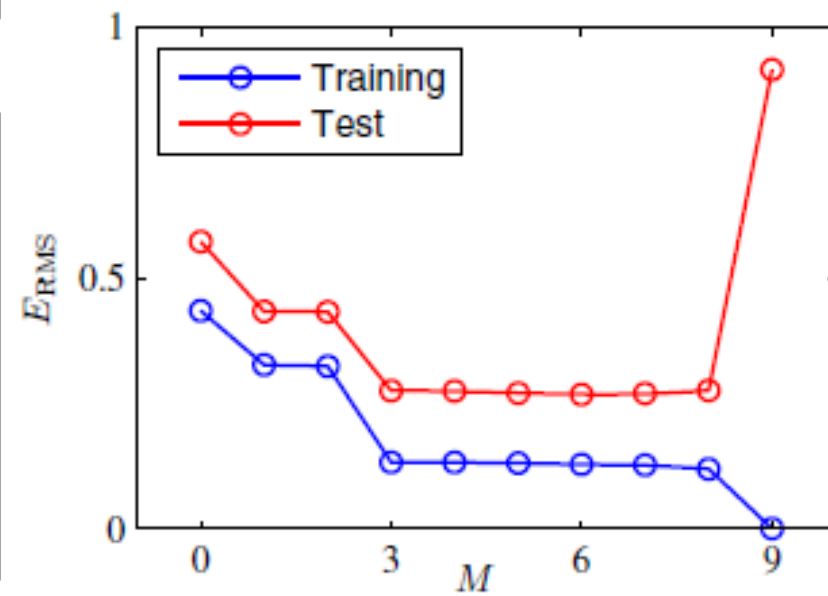
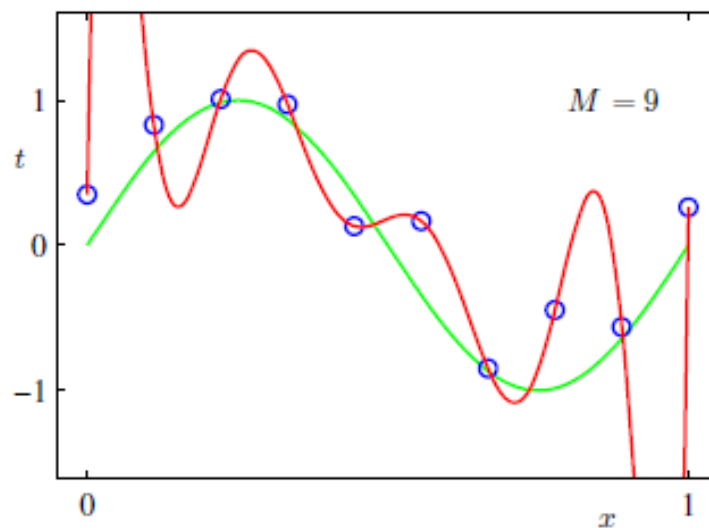
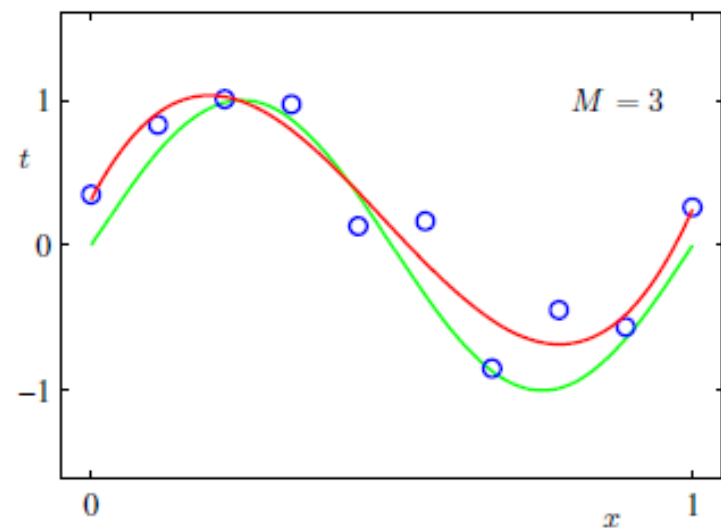
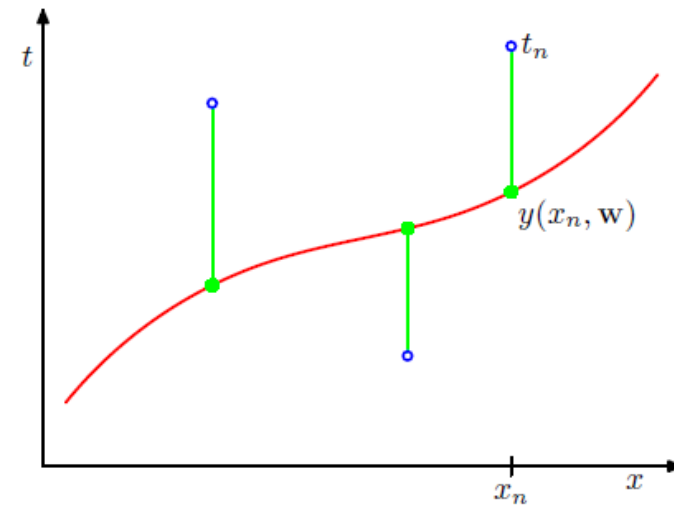
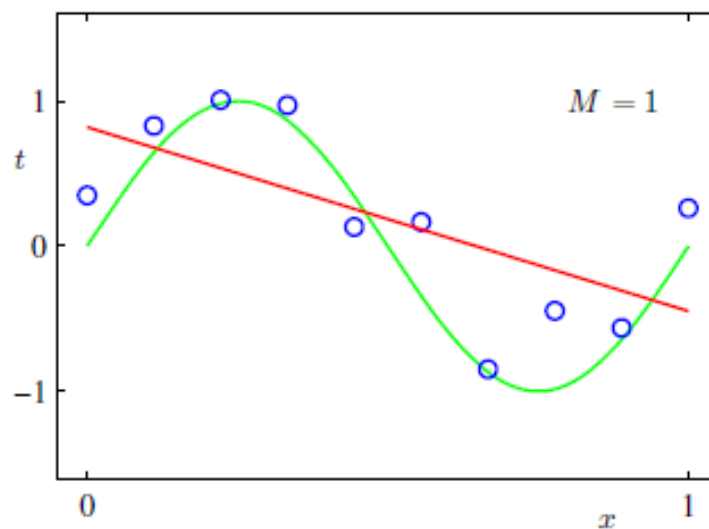
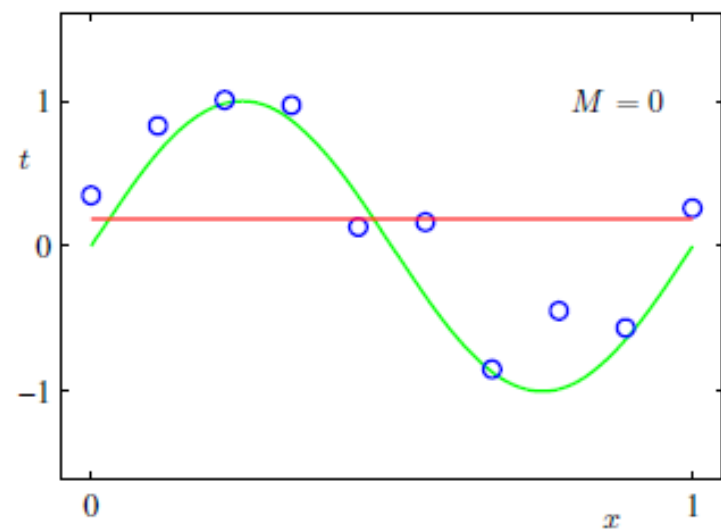
$x$	$y$	$\varepsilon$
Instances		



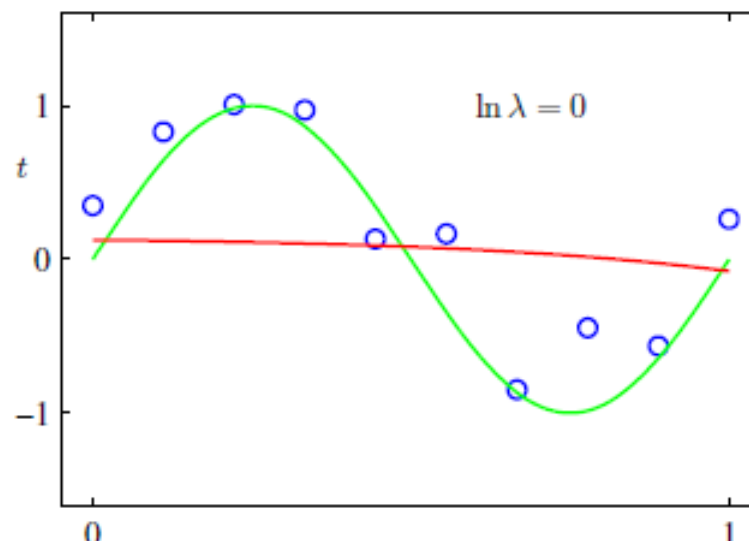
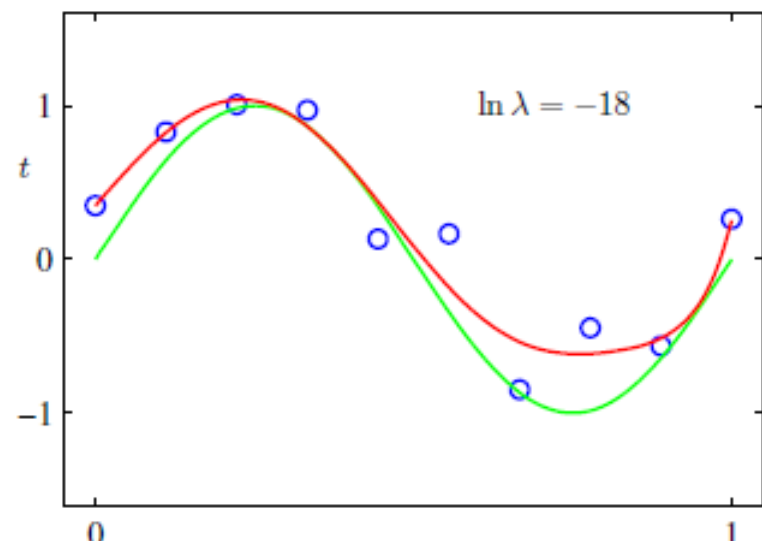
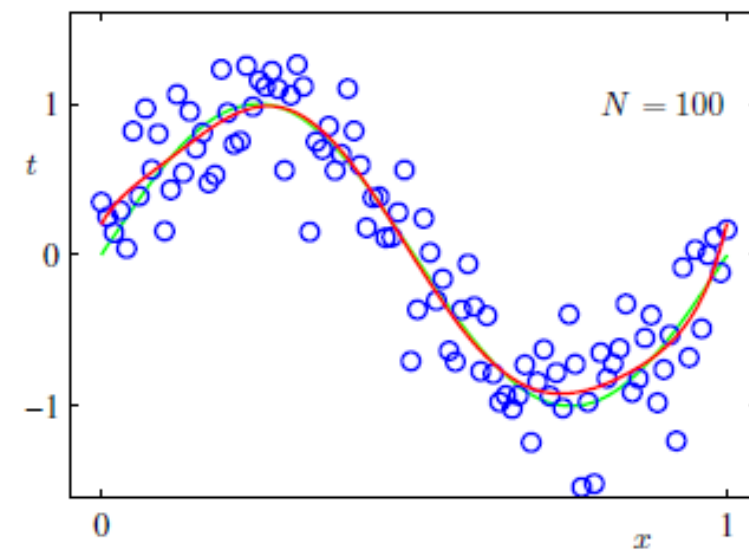
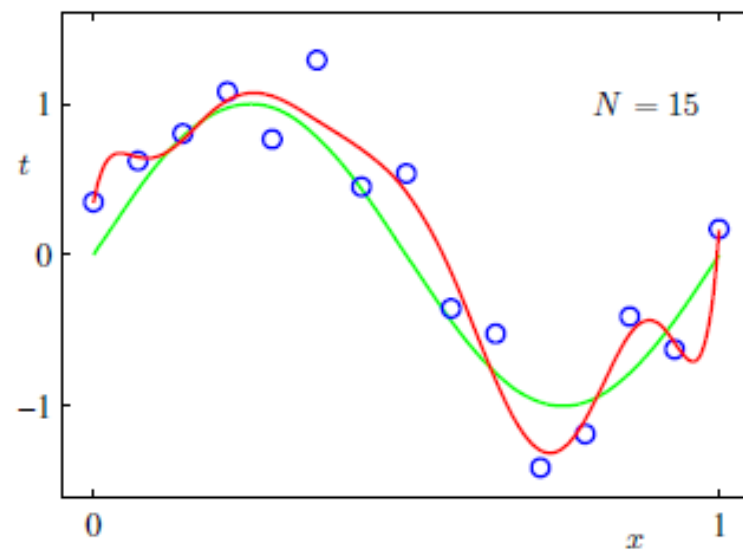
# Machine learning. Introduction



# Model selection

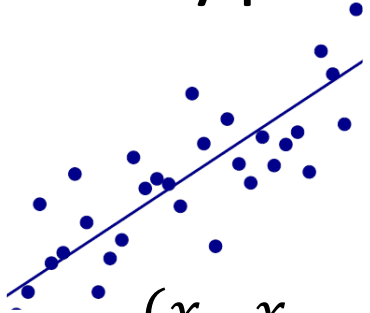


# Over-fitting?



# Concepts review

# Type of problems, data types

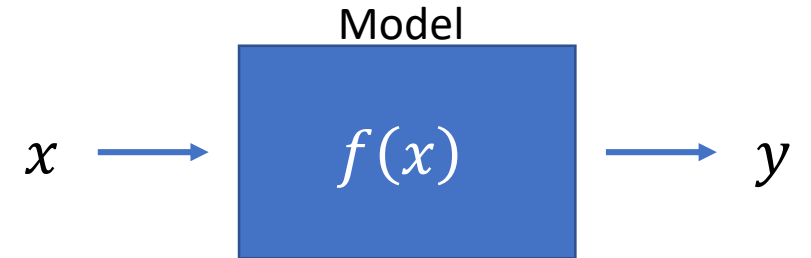


$$x \rightarrow y$$

$$(x_1, x_2, \dots, x_n) \rightarrow y$$

$$(x_1, x_2, \dots, x_n) \rightarrow (y_1, y_2, \dots, y_k)$$

**Supervised**



$$f: X \rightarrow Y$$

**Unsupervised**

ID	$x_1$	...	$x_n$	Category
1	3.532		A	Catx
2	7.234		H	Caty
⋮	⋮		⋮	⋮



ID	Cat y
1	aaa
2	hhh
⋮	⋮

# What is ML

Machine Learning is the science (and art) of programming computers so they can *learn from data*.

Here is a slightly more general definition:

[Machine Learning is the] field of study that gives computers the ability to learn without being explicitly programmed.

—Arthur Samuel, 1959

And a more engineering-oriented one:

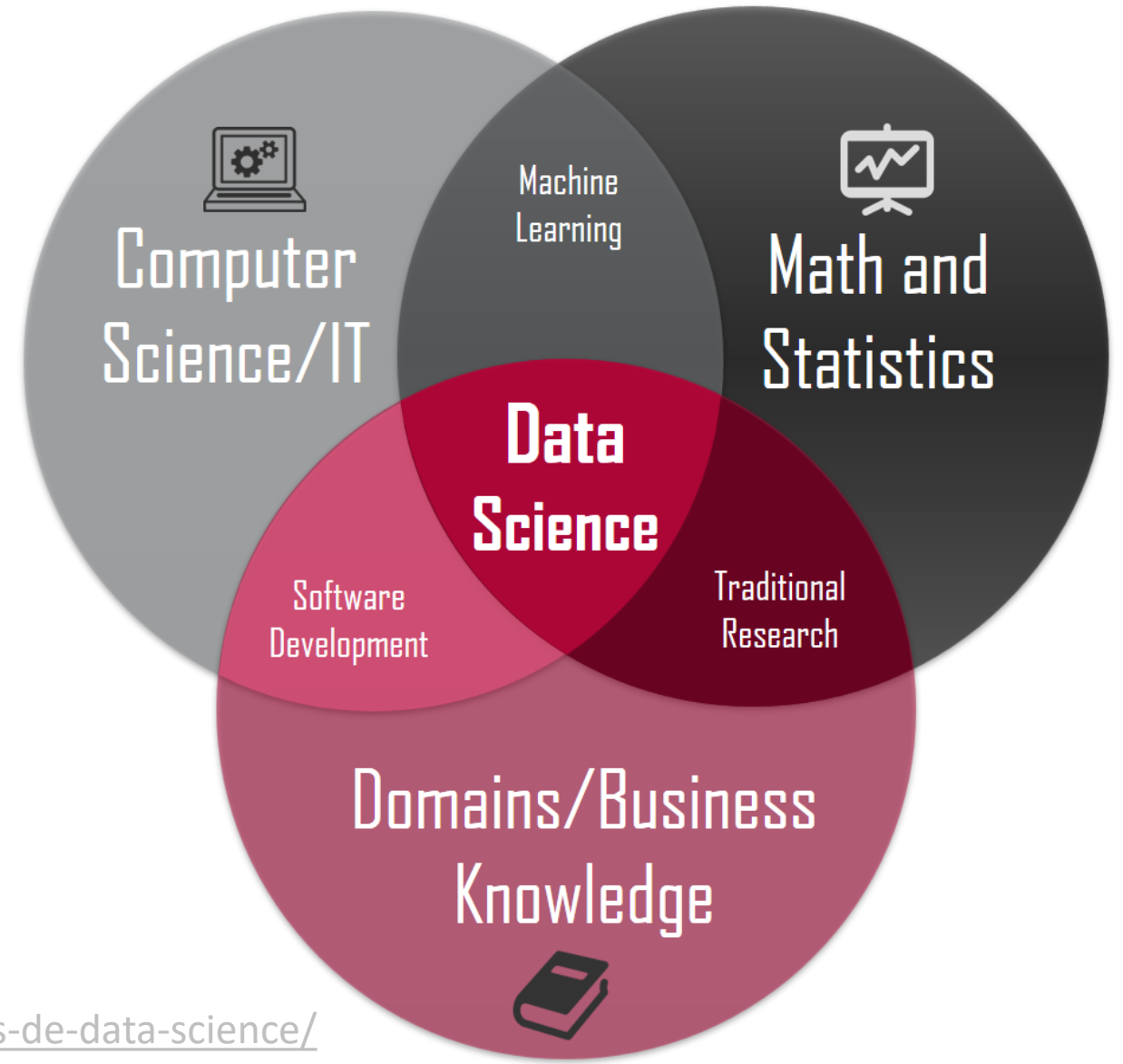
A computer program is said to learn from experience  $E$  with respect to some task  $T$  and some performance measure  $P$ , if its performance on  $T$ , as measured by  $P$ , improves with experience  $E$ .

—Tom Mitchell, 1997



# Related areas. Data Science

- Data manipulation (50 – 80 %)
- Machine Learning
- Deep learning



# Applications in this course

Hydrology, geothermal, radioactive waste disposal, oil and gas, CO2 sequestration, mining

- Fluid volume production from a well
- Porosity from depth (Pyrch)
- ...

Mostly based on

- <https://github.com/GeostatsGuy>
- Guangren Shi, 2014. Data Mining and Knowledge Discovery for Geoscientists

# Types of Machine learning algorithms

- Supervised

- k-Nearest Neighbors
- Linear Regression
- Logistic Regression
- Support Vector Machines (SVMs)
- Decision Trees, Ensemble methods
- Neural networks

- Unsupervised

- Clustering: K-means, Hierarchical Cluster Analysis (HCA)
- Visualization and dimensionality reduction (Kernel)? PCA, t-distributed Stochastic Neighbor Embedding (t-SNE)

- Reinforcement learning

- Batch and Online learning

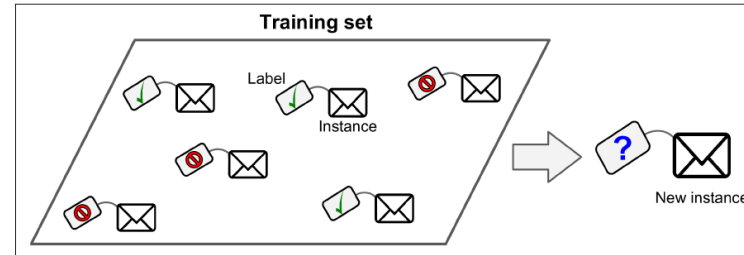


Figure 1-5. A labeled training set for supervised learning (e.g., spam classification)

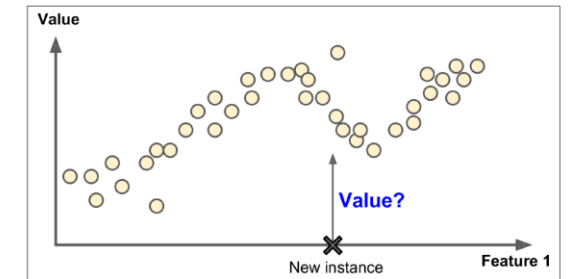


Figure 1-6. Regression

# Grading

- 20% Theoretical
- 20% Computational exercises
- 20% Exams
- 20% Oral presentation of application of ML in Geosciences
- 20% MOOC (Coursera, Udemy, ...)

# Current research (Journals)

- Mathematical Geosciences
  - [Neural Networks](#)
  - [Machine learning](#)
- Computer and Geosciences
  - [Neural networks](#)
  - [TensorFlow](#)
  - [Machine learning](#)
- Computational Geosciences
  - [Neural Networks](#)
  - [TensorFlow](#)
  - [Machine learning](#)

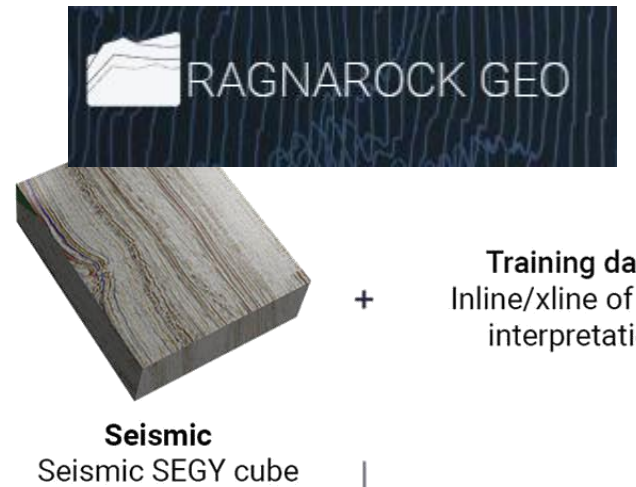
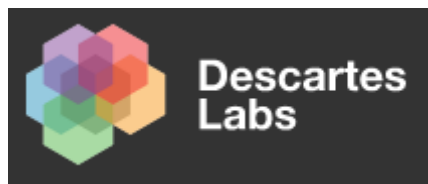
# Table of Content (TOC)

Índice Temático			
Unidad	Tema	Horas	
		Teóricas	Prácticas
1	Panorama general sobre machine learning	2	0
2	Proyecto aplicado de machine learning	4	4
3	Modelos lineales y regresión logística	3	3
4	Máquinas de soporte vectorial (Support Vector Machines)	3	3
5	Métodos basados en árboles de decisión	6	6
6	Modelos basados en teoría de gráficas	3	3
7	Aprendizaje no supervisado	3	3
8	Reducción de la dimensionalidad	3	3
9	Redes neuronales y aprendizaje profundo	6	6
Total de horas:			
Suma total de horas:		64	

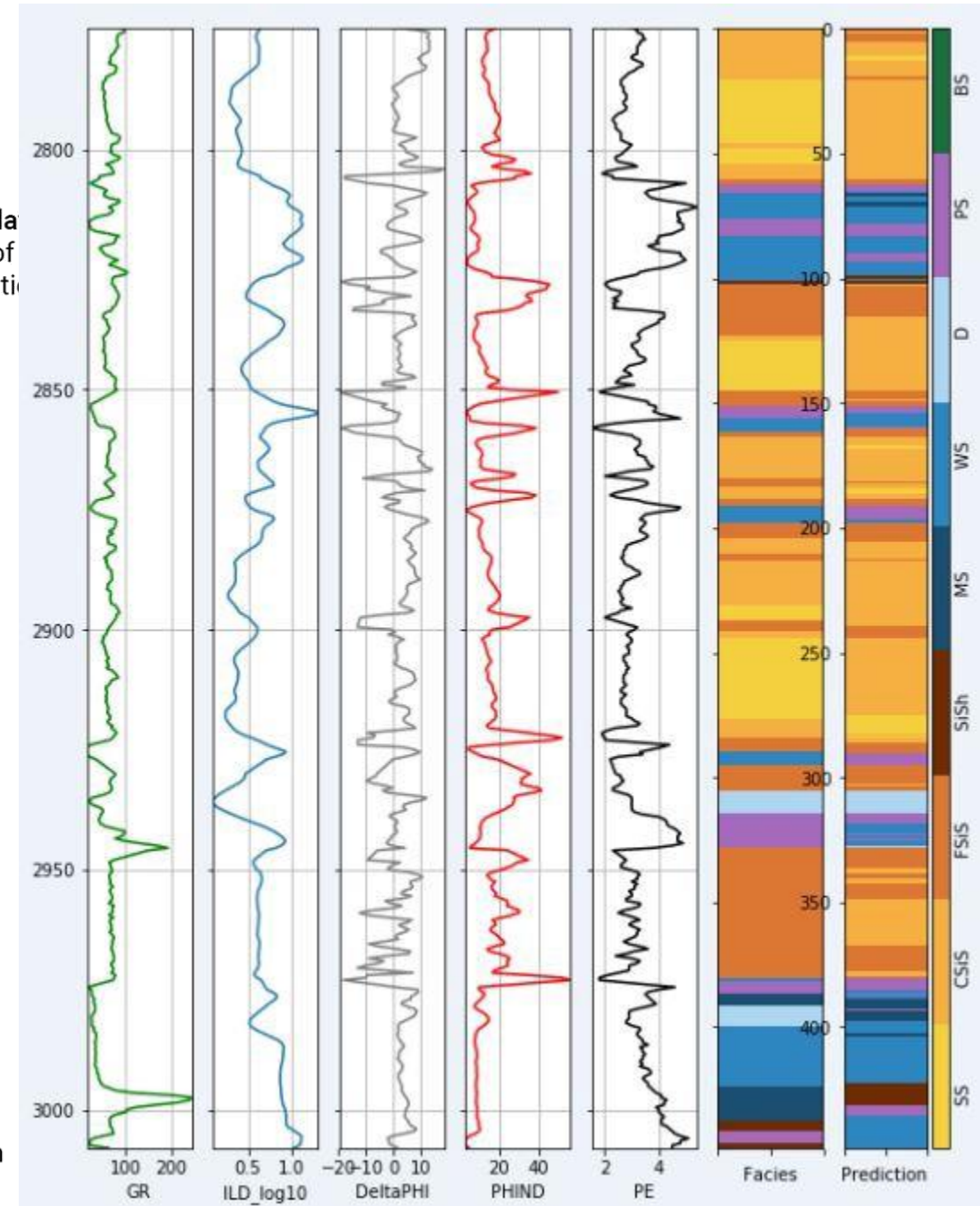
# Bibliography

# Companies

- All big companies +
- Well spacing
- Earth models
- Seismic horizons interpretation



**Output**  
3D model in your seismic interpretation software with your interpretation throughout the cube



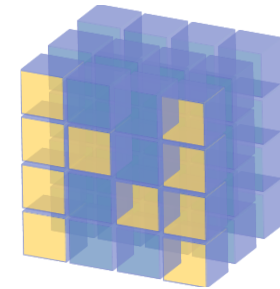


# Concepts review

# Software stack



Visual Studio Code



NumPy

# See also

- [YouTube Michael Pyrcz 00 Machine Learning: Introduction](#)
- [YouTube Michael Pyrcz 06 Machine Learning: Intro to Machine Learning](#)

# Machine Learning for Geosciences

By Francisco Mendoza

[mentofran@gmail.com](mailto:mentofran@gmail.com)



Course material:

<https://github.com/mathphysmx/teaching-ml>