

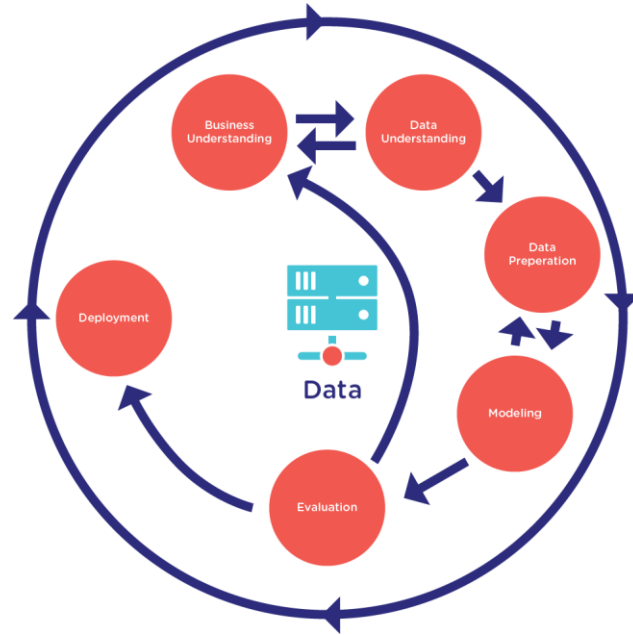
Auxiliar 1

Introducción al curso y Python

Camilo Escobar, Diego Garrido, Sebastián Santana.

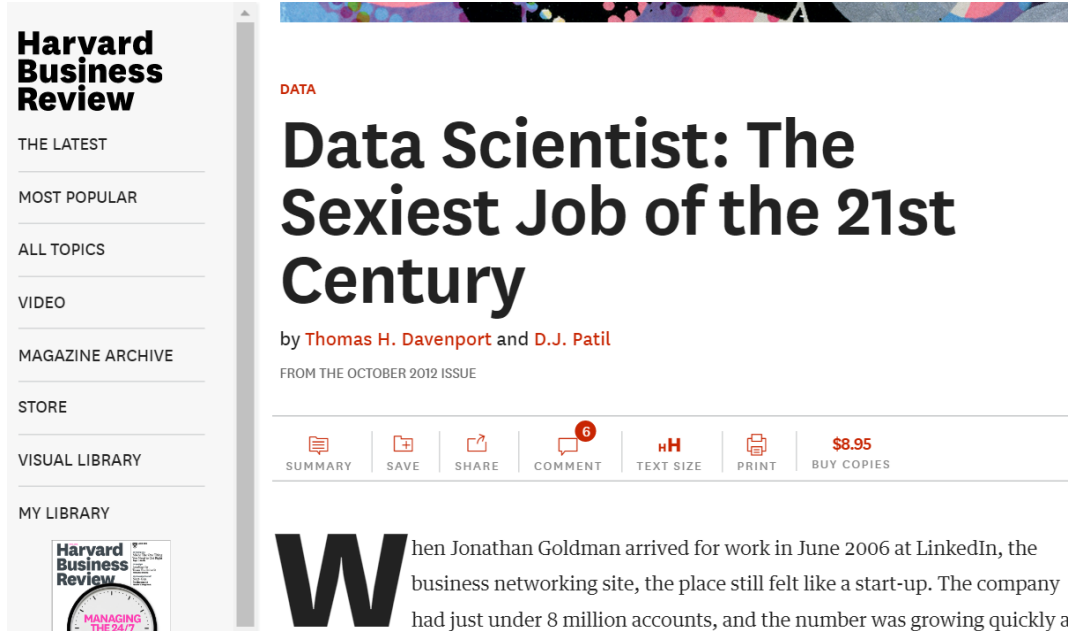
Qué hace Data Science?

Data Science tiene la misión de modelar, analizar, entender, visualizar y extraer conocimiento a partir de datos.



Quién hace Data Science?

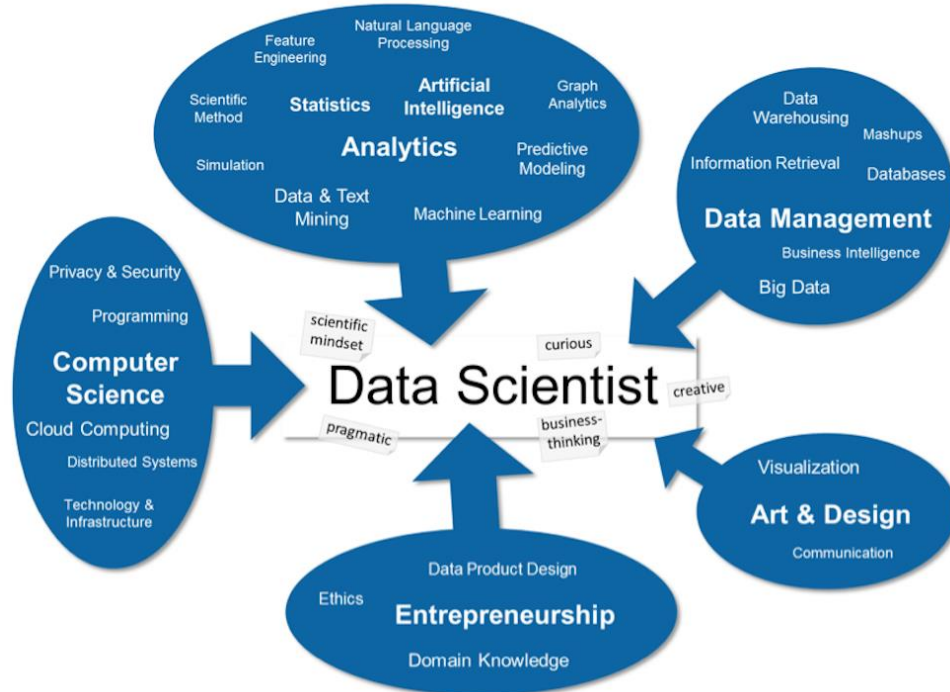
Data Scientists son los profesionales de Data Science, que necesitan conocer mas estadística que un ingeniero de software, y saber mas ingeniería de software que un estadístico



The screenshot shows the Harvard Business Review website. On the left is a navigation menu with links: THE LATEST, MOST POPULAR, ALL TOPICS, VIDEO, MAGAZINE ARCHIVE, STORE, VISUAL LIBRARY, and MY LIBRARY. Below the menu is a small image of the magazine cover titled 'MANAGING THE 24/7'. The main content area features the article 'Data Scientist: The Sexiest Job of the 21st Century' by Thomas H. Davenport and D.J. Patil, from the October 2012 issue. Below the title is a row of interactive icons: SUMMARY, SAVE, SHARE, COMMENT (with a red notification bubble containing the number 6), TEXT SIZE (with an 'H' icon), PRINT, and BUY COPIES (\$8.95). The article text begins with a large 'W' and the sentence: 'hen Jonathan Goldman arrived for work in June 2006 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as'.

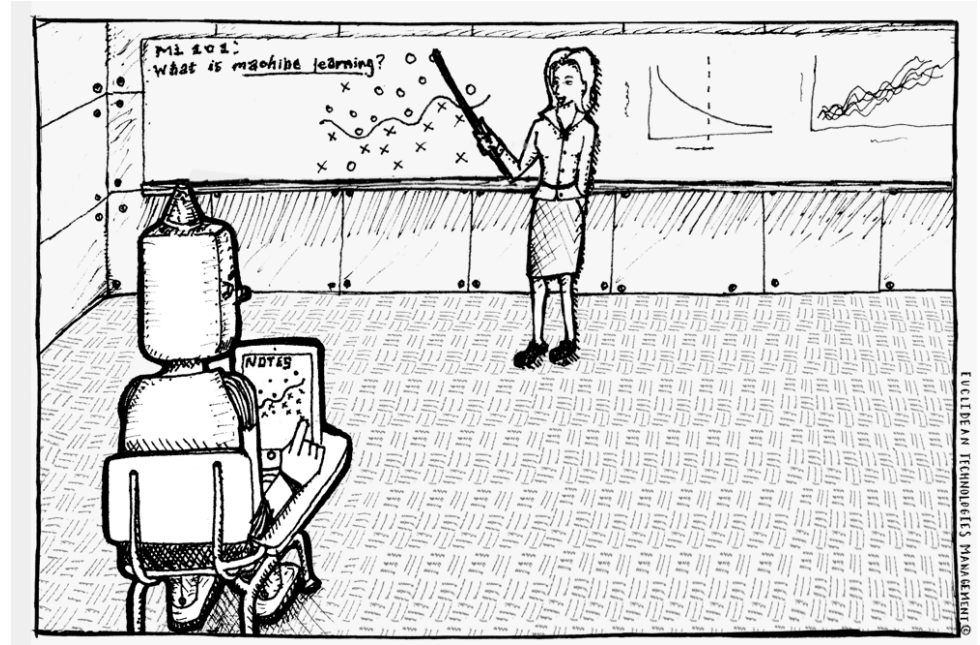
Habilidades interdisciplinarias

Un Data Scientist posee conocimientos (y creatividad!) en modelación, visualización, bases de datos y programación



Machine Learning

Machine Learning es un área cuyo objetivo es desarrollar **algoritmos** que permitan a las **computadoras aprender**.





Aprendizaje De Máquinas

■ Aprendizaje Supervisado

Predictivo: Aprender un mapping de inputs x a outputs y .

Tipos: Regresión, Clasificación

■ Aprendizaje No Supervisado:

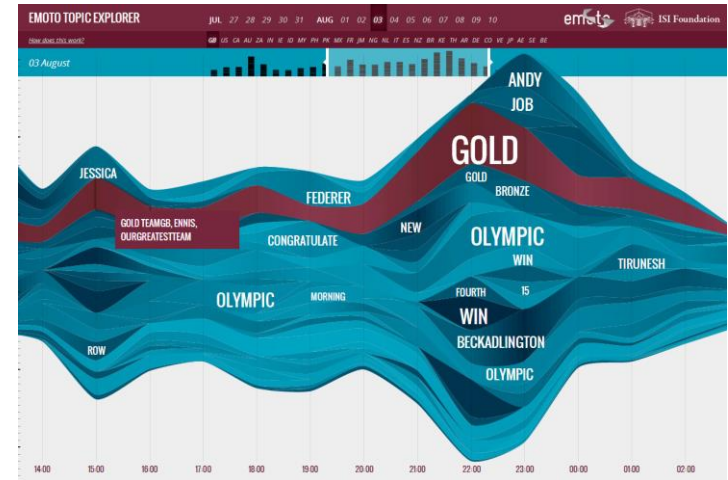
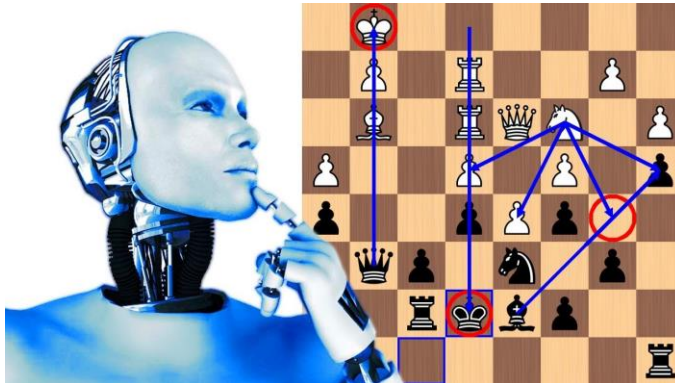
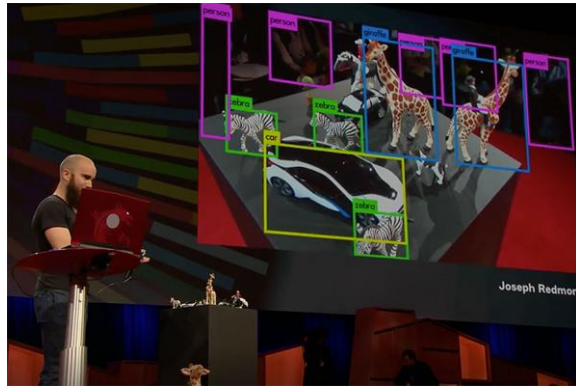
Descriptivo: Encontrar patrones interesantes a partir de inputs x .

Tipos: Discovering cluster, latent factors, graph structure, matrix completion

■ Aprendizaje Reforzado

Recompensa y castigo

Más Ejemplos





Python cuenta con diversas implementaciones, distribuciones, herramientas y gestores de paquetes.

1. Pip: Python Package Index.
2. Anaconda: Python distribution/installer.
3. Conda: Binary Package Manager.
4. Jupyter: Web-Based Interactive Notebook.
5. Django: Framework Web.
6. CPython: C bytecode interpreter.
7. Jython: Java bytecode interpreter.
8. IronPython: .NET bytecode interpreter.
9. PyPy: Just-in-time compiler (RPython to machine code).
10. MicroPython: Microcontroller compiler.

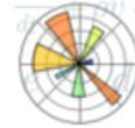
Python dispone de un rico ecosistema compuesto de librerías open-source para matemáticas, estadística, machine learning y ciencia en general.



NumPy
Base N-dimensional
array package



SciPy library
Fundamental library
for scientific
computing



Matplotlib
Comprehensive 2D
Plotting



IPython
Enhanced
Interactive Console



Sympy
Symbolic
mathematics



pandas
Data structures &
analysis



Qué aprenderemos de Python:

1. Programación: Ecosistema, sintaxis, librerías, paradigmas.
2. Manipulación: Numpy, Pandas.
3. Visualización : Matplotlib, Seaborn.
4. Modelación: Scikit-Learn, Gensim, Tensorflow/Keras.

Evaluaciones

Laboratorios

1. 7 laboratorios en clase auxiliar, se elimina el peor.
2. Grupos de máximo 3
3. Laboratorio individual si no asisten
4. Empiezan el 28 de Septiembre
5. Plazo 6, días, se entrega el Jueves de la semana siguiente antes de las 23:59.

Proyecto Semestral

1. 3 Presentaciones (10%, 20%, 30%)
2. 1 Informe final (40%)
3. Grupos de máximo 3
4. Proyecto con empresa.
5. Pregunta interesante que se pueda responder con datos

Otras evaluaciones

CTP

1. Sobre lecturas (5), presentaciones (3) y taller de ética (1).

2. Se eliminarán los peores 2.

Examen

1. Fecha a definir, semana 15 o en semana de exámenes.

Evaluaciones

Para aprobar el curso se necesita lo siguiente:

1. Nota control ≥ 4.0
2. Nota promedio laboratorios ≥ 4.0
3. Nota proyecto semestral ≥ 4.0
4. Nota CTP ≥ 5.0

La nota final del curso se determina de la siguiente manera:

1. Nota control 20%
2. Nota promedio laboratorios 40%
3. Nota proyecto semestral 20%.
4. Nota CTP 20%