

Python: Machine Learning, Optimización y Aplicaciones (III edición)

Dr. Daniel Gutiérrez Reina, dgutierrezreina@us.es
Dr. Sergio Toral Marín, storal@us.es

D. Juan Pedro Pérez Alcantara
Dr. Manuel Perales Esteve
D. Jaime Martel Romero-Valdespino
D. Mario Rivas Sánchez



Módulos del Curso

- ▶ Módulo 1: Conocimientos Básicos de Python y sus Módulos Principales. (20 horas).
- ▶ Módulo 2: Machine Learning en Python: Regresión, Clasificadores y Clustering. (16 horas).
- ▶ Módulo 3: Técnicas de Optimización en Python. (16 horas).
- ▶ Módulo 4: Deep Learning con TensorFlow y Keras. (16 horas).
- ▶ Módulo 5: Aplicaciones. (16 horas).



Fechas importantes

- ▶ Preinscripción: Desde el 1 Junio de 2018 (también abierta durante el plazo de matriculación).
- ▶ Matrícula: Del 1 al 20 de Septiembre de 2018.
- ▶ Inicio-fin curso: 16 de octubre a 4 de diciembre de 2019.



Horarios

- ▶ Módulo 1: 16/10/2019 - 24/10/2019, Miércoles, Jueves y Viernes de 17:00 - 21:00 horas.
- ▶ Módulo 2: 25/10/2019 - 06/11/2019, Miércoles, Jueves y Viernes de 17:00 - 21:00 horas.
- ▶ Módulo 3: 07/11/2019 - 14/11/2019, Miércoles, Jueves y Viernes de 17:00 - 21:00 horas.
- ▶ Módulo 4: 15/11/2019 - 22/11/2019, Miércoles, Jueves y Viernes de 16:00 - 20:00 horas.
- ▶ Módulo 5: 27/11/2019 - 04/12/2019, Miércoles, Jueves y Viernes de 16:00 - 20:00 horas.



Módulo 1

- ▶ Conceptos básicos de programación en Python.
- ▶ Módulo numpy.
- ▶ Módulo matplotlib.
- ▶ Módulo pandas.
- ▶ Módulo Scipy

Conocimientos Básicos de Python y sus Módulos Principales



Módulo 2

► Regresiones: Regresión lineal simple y múltiple, errores en la estimación y overfitting, regresión Ridge y Lasso, aproximaciones no paramétricas.

► Clasificadores: Introducción, clasificadores lineales (regresión logística), overfitting, árboles de decisión, ensamble de clasificadores (boosting), métricas de clasificación, aproximaciones Big Data.

► Clustering: k-means

Machine Learning en Python: Regresión,
Clasificadores y Clustering



Módulo 3

► Introducción a los métodos de optimización meta heurísticos: Métodos de búsqueda local basados en trayectorias. Métodos de búsqueda global basados en poblaciones. Algoritmos basados en enjambre. Programación genética.

► Introducción al módulo de optimización DEAP: Optimización de problemas combinatorios. Optimización de problemas con variables continuas. Optimización multi-objetivo. Optimización de problemas con variables continuas con PSO. Programación genética.

► Modelado de un problema desde cero.

Técnicas de Optimización en Python



DISTRIBUTED
EVOLUTIONARY
ALGORITHMS IN
PYTHON

Módulo 4

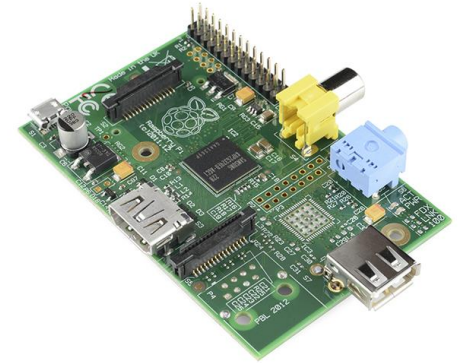
- ▶ Introducción a Deep Learning con TensorFlow y Keras.
- ▶ Redes completamente conectadas: inicialización, regularización y optimizadores
- ▶ Redes neuronales convolucionales
- ▶ Redes neuronales recurrentes

Deep Learning con TensorFlow y Keras

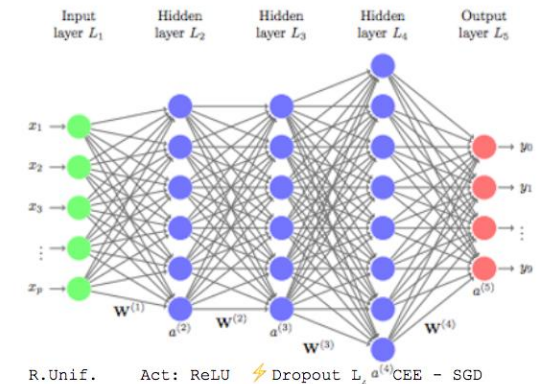


Módulo 5

- Aplicación 1: Análisis de imágenes de satélite
- Aplicación 2: Programación de la Raspberry Pi en Python. APIs y OpenCV
- Aplicación 3: Sistemas de recomendación
- Aplicación 4: Metaheurísticas para la selección de parámetros en redes neuronales



Aplicaciones



Introducción a Python

CONSIDERACIONES PREVIAS:

- ▶ Creado en los años 90s por el holandés **Guido van Rossum**. El nombre se debe al grupo de humoristas *Monty Python*.
- ▶ Python es un lenguaje de **programación interpretado** (Lenguajes interpretados Vs Lenguajes compilados).
 - ▶ C es un lenguaje compilado → Errores en tiempo de compilación.
 - ▶ Python → Los errores saltan en tiempo de ejecución.
- ▶ Lenguaje multiplataforma (Windows, Linux, Mac) → La distribuciones de Linux suelen venir con el interprete de Python ya incorporado.
- ▶ **Open source** (gratis).
- ▶ Está ganando mucha importancia en los último años (diseños web y análisis de datos). **Machine Learning, Big Data, Deep Learning, Artificial Intellingence.**
- ▶ Nosotros vamos a trabajar con la **versión 2.7 de Python** (Aunque hay versiones superiores, ésta es aun la más utilizada). Todas las versiones de Python 2.x son compatibles, hubo un salto con Python 3.x en el que ciertos métodos no son compatibles con Python 2.x.



Introducción a Python

Popularidad de los lenguajes de programación (Publicado en IEEE Spectrum 2018):



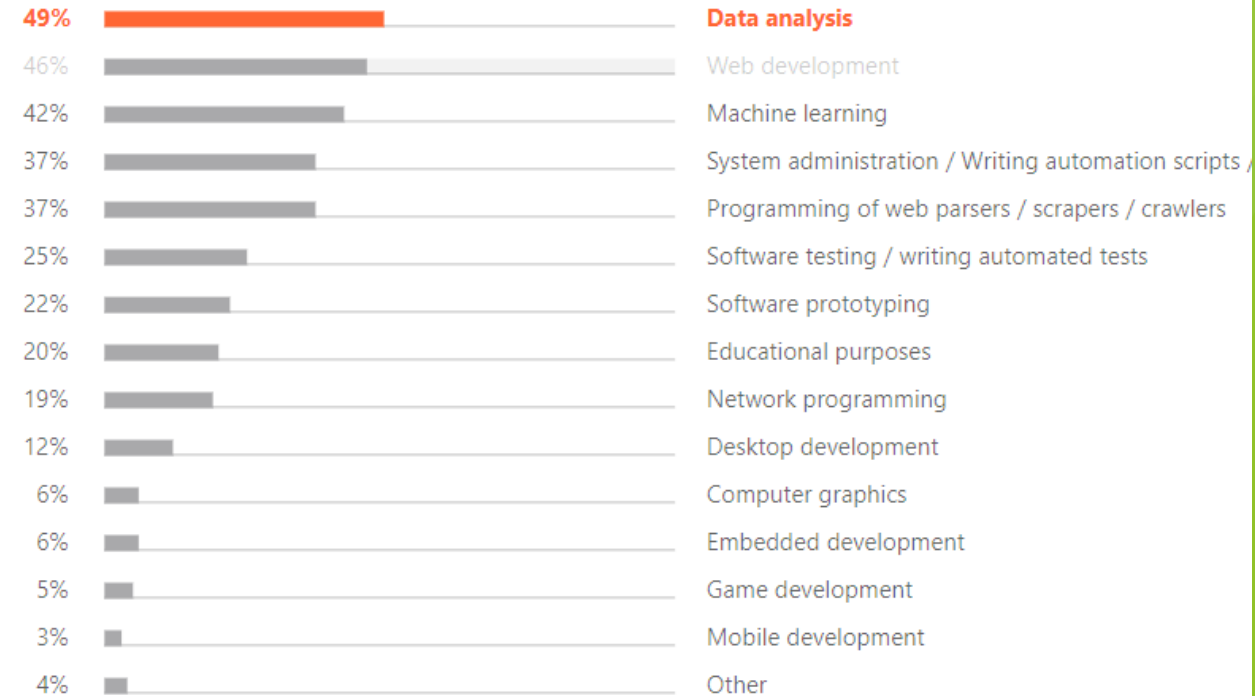
Introducción a Python

Popularity of Programming Language

Worldwide, Sept 2019 compared to a year ago:

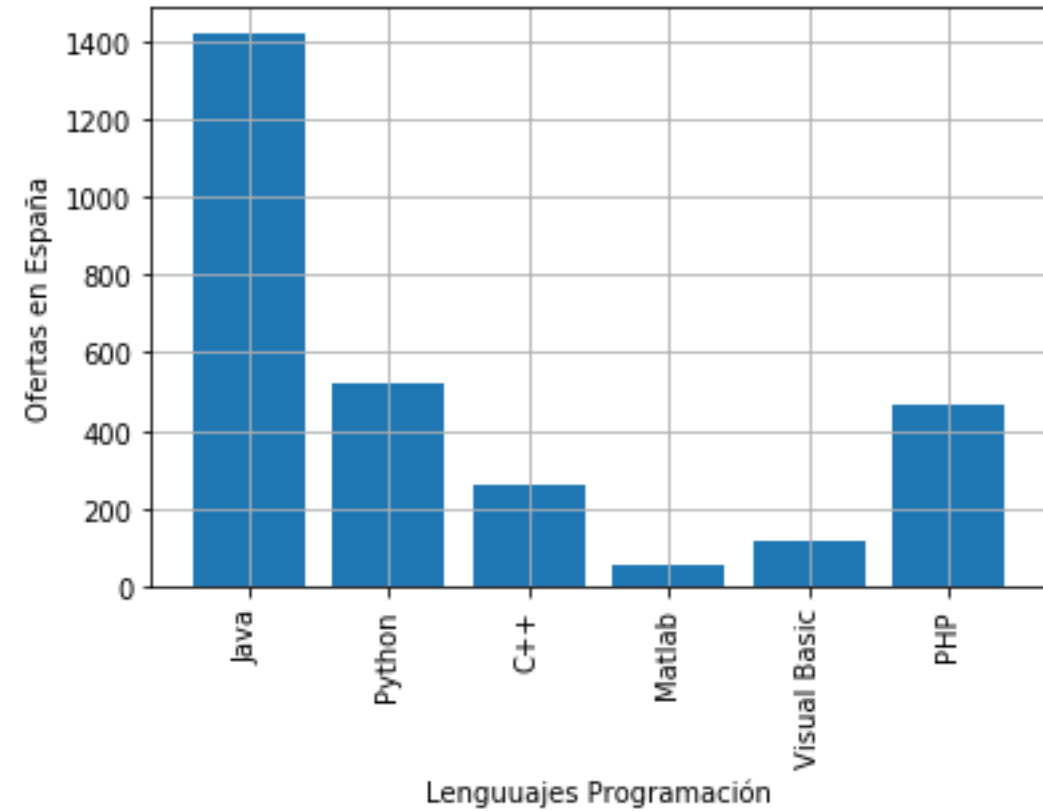
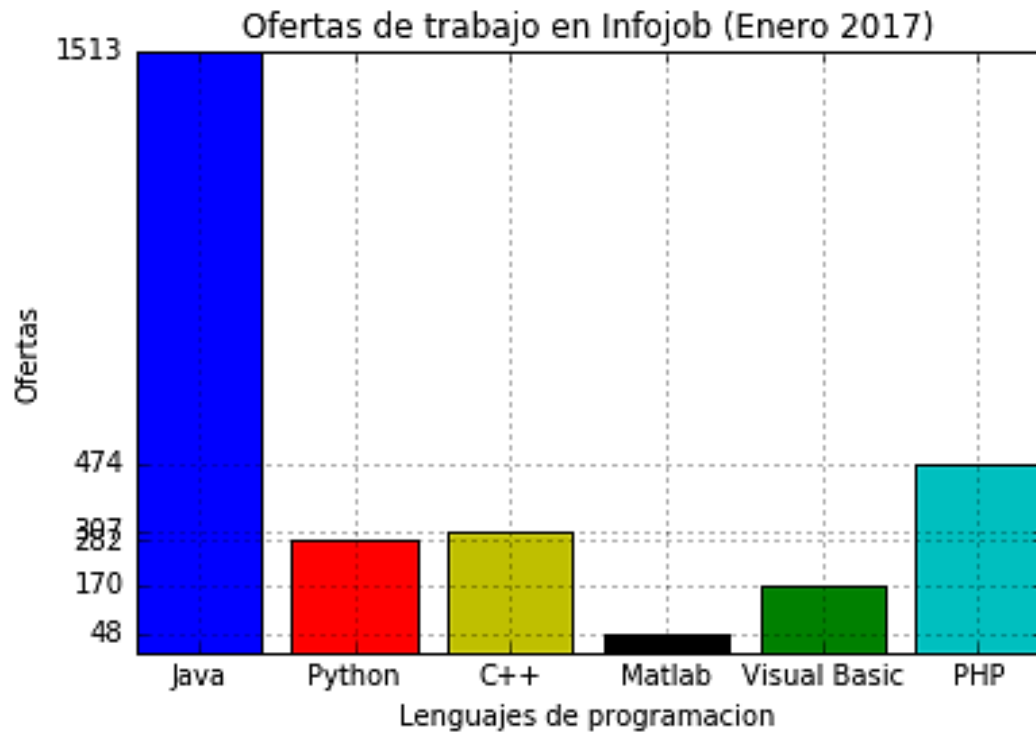
Rank	Change	Language	Share	Trend
1		Python	29.21 %	+4.6 %
2		Java	19.9 %	-2.2 %
3		Javascript	8.39 %	+0.0 %
4		C#	7.23 %	-0.6 %
5		PHP	6.69 %	-1.0 %
6		C/C++	5.8 %	-0.4 %
7		R	3.91 %	-0.2 %
8		Objective-C	2.63 %	-0.7 %
9		Swift	2.46 %	-0.3 %
10		Matlab	1.82 %	-0.2 %

What do you use Python for?



Introducción a Python

Ofertas de trabajo en Infojobs:



Optimización

Conceptos básicos: ¿Qué es optimizar?



Optimización

Conceptos básicos: ¿Cómo resolver problemas de optimización?

- ▶ Fuerza bruta: Evaluar todas las posibles soluciones. Poco eficiente computacionalmente o imposible. Muy ineficiente y poco práctico.
- ▶ Búsqueda aleatoria: Búsqueda aleatoria sin ninguna guía. No garantiza soluciones óptimas, ni quasi-óptimas. Simplemente reduzco el número de evaluaciones.
- ▶ Técnicas basadas en gradiente: Requiere la existencia de derivadas; búsqueda local- problema de óptimos locales.
- ▶ Algoritmos meta heurísticos: Tenemos que ser capaces de evaluar las soluciones. Medir la calidad. No hay garantías de obtener el óptimo.



Optimización

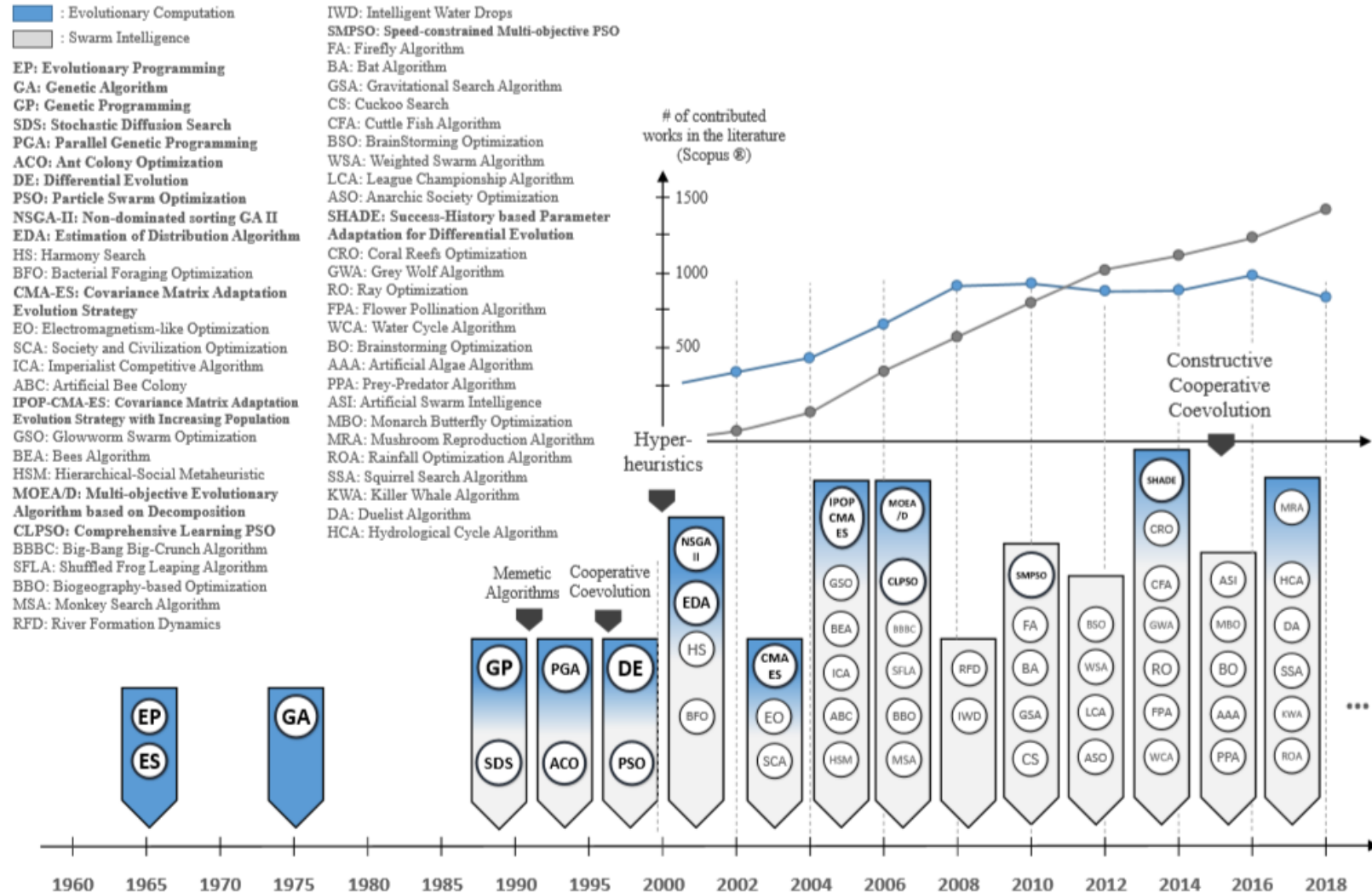
Algoritmos metaheurísticos

- ▶ Son algoritmos estocásticos (probabilístico) que se aplican a problemas tipo “I know it when I see it”.
- ▶ Normalmente no tengo un modelo matemático que quiero optimizar. Es decir, se poco sobre el problema.
- ▶ Pero sí puedo evaluar el problema → Simulaciones.
- ▶ El coste de evaluar el problema es bajo. Lo tengo que hacer muchas veces ☹.



Optimización

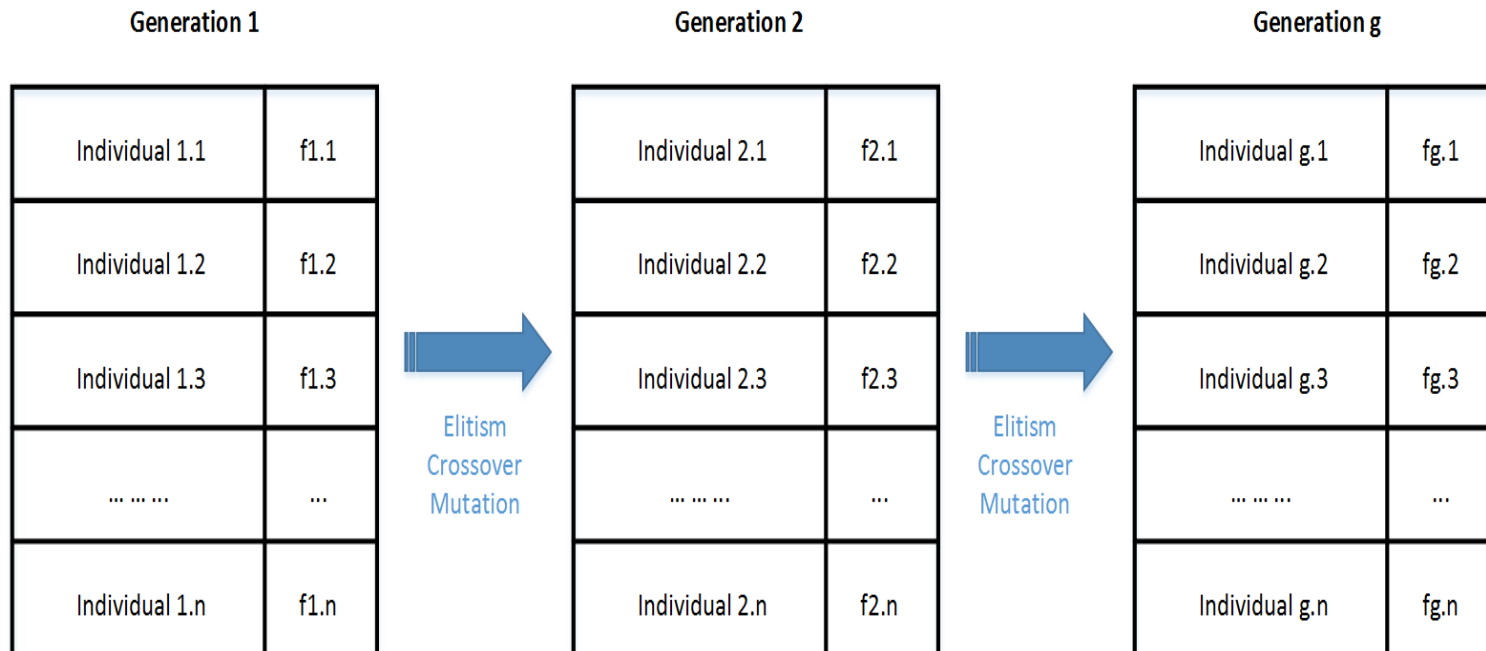
Evolución:



Optimización

Idea global:

Población de individuos o posible soluciones que evolucionan a lo largo de un número de generaciones, creándose mejores individuos mediante operaciones genéticas (cruce y mutación).



El problema del viajero (Traveling Salesman Problem)

Objetivo: recorrer todas las ciudades minimizando la distancia recorrida.

Es un problema NP duro. Debemos probar todas las posibles combinaciones!!!



Número de combinaciones??

$(N-1)! / 2 \rightarrow$ Si consideramos que la $d_{ij} = d_{ji}$

$N= 5 \rightarrow 60$

$N= 10 \rightarrow 1.814.400$

$N= 15 \rightarrow 653.837.184.000$

...

El problema del viajero (Traveling Saleman Problem)

Vamos a utilizar Google colab:

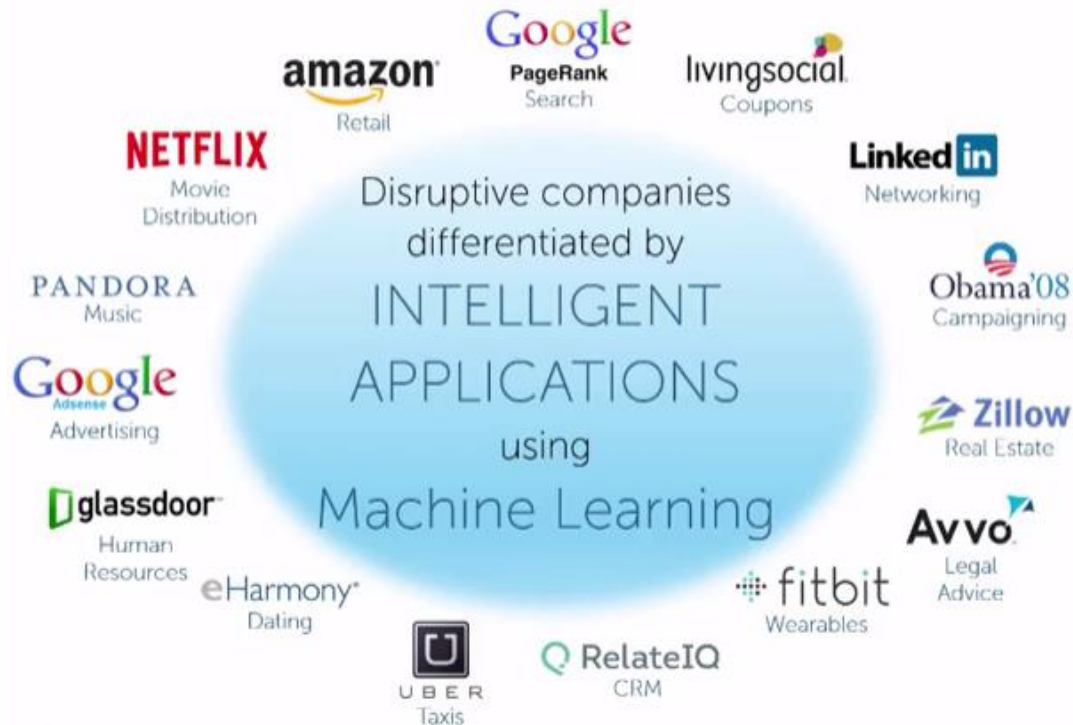
<https://colab.research.google.com/notebooks/welcome.ipynb#recent=true>

Respositorio de Github:

https://github.com/Dany503/CFP_2019.git

Archivo → Abrir cuaderno → Pestaña Github

Machine Learning y Deep Learning



► Sobreabundancia de recursos e información:

- Sistemas de recomendación, películas, música, productos diversos
- Conexión de personas, servicios en tiempo real, detección de fraudes, detección de spam, publicidad dirigida, fijación dinámica de precios



Machine Learning - ejemplo I

► Análisis de sentimiento sobre opiniones online

Sort by: Filter by:

Showing 1-10 of 559 reviews (Verified Purchases). [See all 638 reviews](#)

★☆☆☆☆ Expired products when arrived.

By [Amazon Customer](#) on July 19, 2016

Flavor: DHA and Probiotic Rice | Size: 8 Ounce (Pack of 6) | [Verified Purchase](#)

I bought 2 6 pack cartons and all were expired 6 months before my order. Very disappointed. Order was placed June of 2016.



► [Comment](#) | 35 people found this helpful. Was this review helpful to you? [Report abuse](#)

★★★★★ My Son loves it...

By [Daisyflower](#) on June 17, 2014

Flavor: Organic Brown Rice | Size: 8 Ounce (Pack of 6) | [Verified Purchase](#)

My son loves this and cant seem to have enough of this. I always mix it with my breast milk. He loves it even with formula. I called gerber and checked about the level of arsenic present in it (Call me paranoid but I need to sort certain things before I give it my baby) So they they use California rice which is naturally low in arsenic and the levels that are found is far less than the ones specified by FDA. Hope this helps all those nervous mamas out there...

► [Comment](#) | 119 people found this helpful. Was this review helpful to you? [Report abuse](#)

The manufacturer commented on the review below:

★★★★★ Great product!

By [wrm](#) on June 1, 2016

Flavor: Oatmeal | Size: 8 Ounce (Pack of 6) | [Verified Purchase](#)





When our son was about 4 months old, our doctor said that we could start trying to give him cereal. After reading up on the matter, we decided to go with oatmeal cereal

► Predecir la orientación semántica de los contenidos, entrenando un clasificador a partir del rating de las opiniones



Machine Learning - ejemplo I

- Objetivo del clasificador:
 - Clasificar reviews como positivas o negativas
- Aprendizaje:
 - Reviews anotadas
- Métricas: precisión y recall

		Predicción	
		 $\hat{y}_i = +1$	 $\hat{y}_i = -1$
Valor real	 $y_i = +1$	Verdaderos positivos	Falsos negativos
	 $y_i = -1$	Falsos positivos	Verdaderos negativos

recall

precision

