



Aprende Machine Learning

antes de que sea demasiado tarde

GENERAL

¿Machine Learning en la
Nube? Google Colaboratory
con GPU!

2019-01-10 10:00

Por increíble que parezca, ahora mismo tenemos disponible una cuenta gratuita para programar nuestros modelos de Machine Learning en la nube, con Python, Jupyter Notebooks de manera remota y hasta con GPU para poder aumentar nuestro poder de procesamiento.... gratis! sí sí... esto no es un «cuento del tío» ni tiene ninguna trampa!... Descubre cómo aprovecharlo en este artículo!

Machine Learning desde el Navegador

Primero lo primero. **¿Porqué voy a querer tener mi código en la nube?** Pues bien, lo normal (¿ideal?) es que tengamos un **entorno de desarrollo local** en nuestro propio ordenador, un entorno de pruebas en algún servidor, staging y producción. Pero... ¿qué pasa si aún no tenemos instalado el ambiente?, o tenemos conflictos con algún archivo/librería, versión de Python... ó por lo que sea no tenemos espacio en disco... ó hasta si nos va muy lento y no disponemos en -el corto plazo- de mayor procesador/ram? O hasta por simple comodidad, está siempre bien tener a mano una web online, «siempre lista» en donde ya esté prácticamente todo el software que necesitamos instalado. Y ese servicio lo da Google, entre otras opciones. Lo interesante es que **Google Colab** ofrece varias ventajas frente a sus competidores.

interesante es que **Google Colab** ofrece varias ventajas frente a sus competidores.

La GPU.... ¿en casa o en la nube?

¿Una GPU? ¿para que quiero eso si ya tengo como 8 núcleos? La realidad es que para el procesamiento de algoritmos de Aprendizaje Automático (y para videojuegos, ejem!) **la GPU resulta mucho más potente en realizar cálculos** (también en paralelo) por ejemplo las multiplicaciones matriciales... esas que HACEMOS TOooooDO el tiempo al ENTRENAR **nuestros modelos!!!** para hacer el descenso por gradiente ó Toooooo el rato con el **Backpropagation de nuestras redes neuronales...** Esto supone una mejora de hasta 10x en velocidad de procesado... Algoritmos que antes tomaban días y ahora se resuelven en horas. Un avance enorme.

Si tienes una tarjeta Nvidia con GPU ya instalada, felicidades ya tienes el poder! Si no la tienes y no vas a invertir unos cuantos dólares en comprarla, puedes tener toda(*) su potencia desde la nube!

(*)NOTA: Google se reserva el poder limitar el uso de GPU si considera que estás abusando ó utilizando en demasía ese recurso ó para fines indebidos (por ej. minería de bitcoins)

¿Qué es Google Colab?

Google Colab es un servicio en la nube, que nos provee de una Jupyter Notebook a la que podemos acceder con un navegador web sin importar si «en casa» usamos Windows, Linux o Mac. Tiene como grandes ventajas

- Posibilidad de activar una GPU
- Podemos compartir el código fácilmente
- Está basado en jupyter notebook y nos resultará un entorno ya conocido
- Podemos crear libros en Python 2 ó en 3
- Tiene preinstaladas las librerías comunes usadas en datascience y la posibilidad de instalar otras que necesitemos
- Al enlazar con nuestra cuenta de Google Drive, podemos leer desde ahí archivos csv de entrada ó guardar imágenes de salida, etc.

¿Como se usa GoogleColab?

Primero que nada entramos y logueamos con nuestra cuenta de Google en [Colaboratory](#). Ahora ya podemos:

- Crear una nueva notebook:
 - Vamos a «Archivo -> crear nuevo cuaderno en Python 3»
- y habilitar GPU:
 - Vamos a «Entorno de Ejecución -> Cambiar tipo de entorno de ejecución» y elegimos «Acelerador por Hardware» GPU

Enlazar con Google Drive

Una ventaja de enlazar nuestra cuenta con Drive, es porque nos facilita poder subir o descargar archivos. Para subir un archivo seleccionamos «Archivos» del panel izquierdo y damos al botón «subir» como se muestra en la imagen:

Pero si quieres poder usar cualquier archivo, por ej. csv que tengas en tu unidad de drive, deberas ejecutar en una celda:

```
1 from google.colab import drive
2 drive.mount('/content/drive/')

```

Te pedirá que hagas click a un enlace y escribas un código que te dará cuando autorices la

app. Cuando vuelvas y hagas actualizar en el tab de archivos veras tu unidad montada y lista para usar !!

Ejecutar una jupyter notebook de Github

Vamos a abrir una Jupyter Notebook que contiene el ejercicio explicado en el artículo de [Convolutional Neural Networks: Clasificar 70.000 imágenes deportivas](#). Para ello, en el cuadro de «Abrir cuaderno»:

1. seleccionamos GITHUB,
2. copiamos la dirección del repositorio, en nuestro caso <https://github.com/jbagnato/machine-learning/>
3. y le damos a la lupa de buscar.
4. Nos aparecerá el listado con los archivos del repo.
5. Y de allí seleccionamos el cuaderno llamado **Ejercicio_CNN.ipynb**

Veremos que tenemos el mismo Notebook pero en Google Colab 😊

Descargar un recurso al cuaderno

CASI TODO LISTO... pero..... aún **nos queda algo antes de poder ejecutar**. En este ejercicio, necesitamos tener las 70.000 imágenes en sus directorios respectivos.

Para ello, primero descargaremos el ZIP. Creamos una celda nueva y ejecutamos:

```
!wget https://github.com/jbagnato/machine-learning/raw/master/sportimages.zip
```

y veremos que aparece nuestro archivo zip en el listado (dale a «Actualizar» si hace falta)

Descomprimir un archivo en el cuaderno

Y ahora deberemos descomprimirlo, creamos una celda y ejecutamos:

```
!unzip -uq 'sportimages.zip' -d .'
```

RECUERDA habilitar el entorno de ejecución con GPU como vimos antes. Ahora Ya podemos ejecutar todas las celdas y veremos qué rápido ejecuta la CNN con GPU en

podemos ejecutar todas las celdas y veremos que rápido ejecuta la CNN con GPU, en comparación con CPU. Pasa de tardar 4 minutos a sólo 40 segundos.

Instalar otras librerías Python con Pip

Deberemos ejecutar por ejemplo: `!pip install gensim`

Conclusión

Hemos visto que tenemos la opción de tener **nuestro ambiente de desarrollo local** pero también esta alternativa de poder programar, experimentar y trabajar en la nube. Gracias a este servicio podemos tener listo el ambiente en pocos segundos y aprovechar de las ventajas que nos ofrece, sobre todo el uso de GPU que es un recurso del que no todos disponemos. Espero les haya gustado el artículo, no dejen de escribir sus comentarios, dudas y también te animo a que escribas qué temas de Machine Learning te interesa más. Hasta la próxima!

Recuerda tener instalado tu ambiente de desarrollo Local con Anaconda, aquí te explico cómo hacerlo!

Suscripción al blog

Suscríbete al blog y recibe los nuevos artículos directo en tu casilla de correos!

Email:

ENVIAR

NOTA: muchos usuarios reportaron que el email de confirmación y/o posteriores entraron en su carpeta de SPAM. Te sugiero que revises y recomiendo que agregues nuestro remitente a tus contactos.

Nuevo!: **Realiza Pronóstico de ventas con Redes Neuronales con Keras en Python**

Otros Recursos / enlaces

Otros artículos que explican el uso de Google Colab (en inglés):

- [Google Colab Free GPU tutorial](#)
- [Google Colab for Beginners](#)
- [Picking a GPU for Deep Learning](#)

Comparte el artículo:



Relacionado



Instalar ambiente de Desarrollo
Python Anaconda para Aprendizaje
Automático



Tu propio Servicio de Machine
Learning



12 Consejos útiles para aplicar
Machine Learning

desarrollo

Jupyter Notebook

Machine Learning

Python

tutorial

EJEMPLO WEB SCRAPING EN PYTHON: IBEX35® LA BOLSA DE MADRID

PRONÓSTICO DE SERIES TEMPORALES CON REDES NEURONALES EN PYTHON

13 comments



Oscar · febrero 5

Muchas gracias, probé en días pasados con la librería AI GYM que es bastante pesada y funcionó sin problema

Responder



Na8 · febrero 5

Hola Oscar, gracias por comentar tu experiencia! A mi realmente me sorprende tener al servicio la GPU, Memoria y disco... todo gratuito. Saludos!

Responder



Raúl · febrero 5

Gracias por subirlo tan rapido! la verdad es que es una maravilla de herramienta

Responder



Na8 · febrero 6

Hola Raúl, gracias por tu inspiración! jaja! Saludos!

Responder



Riki · febrero 6

Gracias amigo

Responder



Na8 · febrero 6

de nada!, a seguir creando contenidos 😊

Responder



davinchy seiya · marzo 20

hola buen dia , gracias por tus artículos son muy didácticos e interesantes. Tengo duda, le ves alguna desventaja al colab, por otro lado hasta que punto crees se mantendrá gratuito el servicio, o hay algunas opciones de pago?

Responder



Na8 · marzo 21

Hola Davinchy, gracias por escribir!. Inicialmente no le noto desventajas para hacer experimentos y pruebas. Pero no para trabajar, pues si no me equivoco, los ambientes duran 1 día. Es decir, si subes un archivo (por ejemplo un csv) si no me equivoco a las 12 de la noche será borrado y lo tendrás que volver a subir. Lo mismo con el kernel y la memoria que se «resetean».

Imagino que seguirá gratuito varios años, aunque imagino que te refieres a los servicios de «google Cloud» que ofrecen de pago. Por ahora no les afecta porque el mundo del ML y la IA aún son muy nuevos, pero puede que en el futuro si, ó que pongan nuevas limitaciones.

Responder



davinchy seiya · marzo 21

gracias por tu respuesta, muy amable...y si es verdad se resetea y toca volver va subir los archivos, pero en este caso tambien se pude tener los archivos en el drive de Google y así solo toca ejecutar y colocar el código que genera Google y de esta manera no toca cargar todo desde el inico. Muchas gracias y espero que podas seguir subiendo contenido , están didácticos los artículos.Y gracias por tu tiempo.

Responder



Pablo Paredes · marzo 22

Hola Juan, como estas? En primer lugar gracias por compartir tus conocimientos, soy nuevo en esto de ML e IA, tengo una duda con respecto al SET de DATOS, si tengo varias instancias tipificados con fechas y horas de algunos datos, es decir que los atributos de esos datos varían por hora, si quisiera crear una red neuronal a partir de ese conjunto de datos, y analizarlos diariamente, cual seria el mejor esquema como para crear este set?....Un ejemplo seria:

	Velocidad	Capacidad	Performance

22/03/2019 08:00 Dato_A 3 10 0.25
22/03/2019 08:00 Dato_B 4 10 0.15
22/03/2019 09:00 Dato_A 3.5 10 0.35
22/03/2019 08:00 Dato_B 1.5 10 0.1

Aguardo tus comentarios

Responder



Na8 · marzo 24

Hola Pablo, tiene toda la pinta de un problema del tipo «time series». Por lo tanto, te recomiendo que leas los artículos que escribí hace poco, empieza por la primera parte: [Pronóstico de Series Temporales con Redes Neuronales en Python](#).

Espero que te ayude!

Saludos

Responder



Andres · mayo 10

Crack! Todo el blog es una joya!

Muy didáctico.

Gracias!

Responder



Na8 · mayo 16

Hola Andres!, muchas gracias por el comentario! Saludos

Responder

Deja un comentario

Introduce aquí tu comentario...

Visita nuestra Guía de Aprendizaje

Buscar

Search ...

Contacto

Suscripción

Recibe los artículos de Aprende Machine Learning en tu casilla de correo. Cada 15 días y sin Spam!

Email:

ENVIAR

Proudly powered by WordPress | Theme: Eighties by Justin Kopepasah.

Obtén un [navegador compatible](#) para conseguir un reto reCAPTCHA.

¿Por qué tengo que hacer esto?