

Fundamentos de Big Data - Tarea lección 1

Kevin Martínez García

19 de mayo de 2022

Índice

1. Introducción	2
2. Algunos proyectos de Big Data y Machine Learning	2
3. Usuario en <i>Kaggle</i>	2

1. Introducción

En la actividad de la primera lección de Fundamentos de Big Data, se nos solicitó hacer una búsqueda de ejemplos de proyectos que hagan uso del *Big Data* y el *Machine Learning* y que no hubiesen sido vistos en sesiones de teoría. A continuación, exponemos tres proyectos que hacen uso de estas nuevas tecnologías, así como el registro en la plataforma *Kaggle*.

2. Algunos proyectos de Big Data y Machine Learning

Los proyectos que resultan interesantes a comentar, son los que exponemos a continuación.

- **Machine Learning en el LHC:** El gran colisionador de *hadrones* construido por el CERN (Organización Europea para la Investigación Nuclear) hace uso de técnicas de Machine Learning y Big Data para clasificar algunos de los tipos de colisiones que se llevan a cabo. En la actualidad, uno de sus objetivos es el de conseguir acelerar las redes neuronales profundas (*deep neural networks*) que seleccionan y clasifican las colisiones protón-protón. Para conseguirlo, se busca una reducción en la precisión de los parámetros numéricos que caracterizan a la red para reducir el espacio y el tiempo de procesamiento sin sacrificar rendimiento [1].
- **Big Data y Machine Learning en Twitter:** La red social Twitter muestra en su pantalla de inicio una *línea del tiempo*, es decir, una secuencia de las publicaciones de los usuarios a los que seguimos. En la actualidad, Twitter utiliza técnicas de recolección masiva de datos, así como redes neuronales profundas para ofrecer a cada usuario los *tuits* que podrían resultarle más relevantes [2].
- **Siri y el NLP:** *Siri* es el asistente personal en dispositivos de la marca *Apple* y que realiza numerosas funciones para sus usuarios. Este asistente funciona gracias a técnicas de reconocimiento de voz, así como técnicas de procesamiento del lenguaje natural (NLP). Para conseguir que *Siri* reconozca correctamente el discurso de cada usuario, *Apple* utiliza *data-sets* con muestras muy diversas que entrenan a los algoritmos NLP y así, logran reconocer una amplia diversidad de idiomas, acentos y otras particularidades del habla [3].

3. Usuario en *Kaggle*

Cuenta creada en *Kaggle* con nombre de usuario: kevinmg967

Referencias

- [1] CERN. (n.d.). Speeding up machine learning for particle physics. [online] Available at: <https://home.cern/news/news/physics/speeding-machine-learning-particle-physics>.
- [2] Social Media Marketing Management Dashboard. (2021). How the Twitter Algorithm Works in 2022 and How to Make it Work for You. [online] Available at: https://blog.hootsuite.com/twitter-algorithm/#How_the_Twitter_algorithm_works_in_2022 [Accessed 12 May 2022].
- [3] Ilya Dudkin (2018). How Does Siri Work: Technology and Algorithm. [online] Skywell Software. Available at: <https://skywell.software/blog/how-does-siri-work-technology-and-algorithm/>.