

Título del Proyecto: Haciendo Data Science a League of Legends

Nombre del Alumno: Jorge de Andrés González

Titulación: Experto en Big Data

Curso: 2018-2019

Datos: <https://www.kaggle.com/datasnaek/league-of-legends>

Tutores:

- Telefónica: Por Asignar
- UPSA: D. Manuel Martín-Merino Acera

Explicación y Motivación del Trabajo:

En la actualidad, los deportes electrónicos están moviendo grandes cantidades de dinero, llenando estadios y llenando conferencias de gente interesada en los mismos. Algunos ejemplos son el E3 en Los Ángeles, la BlizzCon, o incluso en España tenemos la Madrid Games Week que se celebra en IFEMA y atrae a miles de visitantes.

Dentro de los videojuegos destacan los competitivos profesionalmente, en los que se genera una gran cantidad de dinero y se reparte entre equipos profesionales, que compiten en torneos a diferentes niveles.

De estos juegos competitivos, hay uno que lleva a la cabeza durante varios años: El League of Legends. Este videojuego, en el que Movistar posee uno de los mejores equipos de España, fue inventado por Riot Games en 2009 y desde entonces no ha dejado de crecer, organizando campeonatos mundiales anuales y moviendo millones de euros.

De este modo, mi trabajo versará sobre dicho videojuego. Consistirá en el análisis de partidas reales, siendo más de 51.000 partidas clasificatorias, para poder obtener los elementos intrínsecos que puedan darse en una partida y que puedan afectar en su resultado.

Para ello, se hará un análisis estadístico básico para saber con qué datos se trabaja y en qué rangos se mueven, una limpieza y/o cambios en los datos para estructurarlos correctamente, un análisis exploratorio extenso para obtener todos los detalles de las partidas y como afecta cada elemento, un análisis a través de machine learning para poder predecir el resultado de las partidas mediante diversas técnicas y finalmente unas pequeñas visualizaciones de los datos obtenidos.

Objetivos del Trabajo:

Los principales objetivos de trabajo son los siguientes:

- 1) Obtención de resultados sobre mitos del juego y sobre personajes y utilidades.
- 2) Obtención de estadísticas y demostración de relaciones entre variables que pueden derivar en el resultado final de la partida.
- 3) Estudio, clasificación y predicción mediante diversos métodos de Machine Learning.
- 4) Visualización de algunas estadísticas importantes que indiquen como influyen ciertos parámetros al devenir de las partidas.

Metodología de Trabajo:

La metodología de trabajo consistirá en la subdivisión del mismo en dos partes:

- 1) Parte práctica: Desarrollada principalmente en Python3 en Jupyter Lab. En caso de haber desarrollo en R, se hará en R 3.5.2 bajo el IDE RStudio, siguiendo R-Markdown.
- 2) Parte teórica: Correspondiente a la memoria, donde se desarrollarán teóricamente los métodos seguidos y se expondrán las conclusiones del trabajo.

El trabajo estará controlado por el VCS "GitHub", quedando el proyecto oculto durante su desarrollo a menos que se indique lo contrario, y haciéndolo público el día de la presentación.

Para el control de tareas, se utilizará el tablero Kanban usando GitHub Projects.

Resultados Esperados:

Ante el dataset referido, tras un estudio y comprensión de sus variables, creo que los resultados que se obtendrán serán muy positivos debido a que algunas variables son muy predictoras del resultado final de la partida.

Además, y al contrario que en otros proyectos, en este con 51490 partidas se tiene dataset suficiente para obtener un buen muestreo y aprendizaje de los algoritmos.

Debido a esto, la afinación de los modelos de Machine Learning será el gran reto para conseguir los mejores resultados posibles, que espero estén por encima del 80 u 85% de acierto.