
Proyecto Kaggle

- La presentación y el informe deben ser entregados antes de 20 de Junio 11:59 p.m.
 - La solución puede ser entregada en grupos de máximo 4 personas.
-

1. Descripción del Proyecto

El proyecto, consiste en escoger una competencia en la plataforma Kaggle de las categorías Featured, Masters, Recruiting, Research o Playground por grupo, el objetivo del proyecto es utilizar al menos tres algoritmos correctamente usando los conocimientos adquiridos en el curso para encontrar los mejores modelos. De esta manera, la evaluación del proyecto estará basada tanto en el desarrollo de los algoritmos como en el buen desarrollo de los modelos y la justificación de su uso.

2. Desarrollo del Proyecto

Elección de Competencia

Se debe elegir una competencia de Kaggle de una de las siguientes categorías:

Featured:

Se trata de concursos públicos con premios significativos destinados a resolver problemas comerciales. Los ganadores del premio otorgan al patrocinador una licencia no exclusiva para su trabajo, y presentarán sus resultados a través de una descripción detallada y posible proyección de pantalla.

Masters:

Son concursos abiertos solo para un nivel selecto de Kagglers de élite, o un subconjunto de estos por invitación o criterios de elegibilidad especiales. Estas competencias tienen un valor comercial significativo o datos confidenciales.

Recruiting:

Son competencias públicas en las que los patrocinadores buscan contratar científicos de datos y utilizar la competencia para buscar y evaluar talentos potenciales. No hay equipos, y cada usuario debe exhibir su trabajo individual. Puede competir anónimamente en este tipo de competencia.

Research:

Son competencias públicas en las que los objetivos de la competencia son de naturaleza investigativa / científica o sirven a un bien público. Estas competencias tienden a centrarse en problemas ambiciosos de machine learning a la vanguardia de la tecnología, o problemas con un aspecto significativo de bienestar social. La mayoría de estos concursos incluyen premios en efectivo, aunque los premios también pueden ser invitaciones a conferencias o publicaciones en revistas revisadas por pares. Estas competencias requieren soluciones modelo que se lanzarán como de código abierto.

Playground:

Son competencias públicas configuradas para ser divertidas, extravagantes e impulsadas por las ideas, en lugar de resolver un problema comercial o de investigación específico. Estos son por diversión y no por ningún premio.

Nota: No se puede escoger competencias de la categoría **Getting started**.

El nombre de la competencia escogida se debe publicar en el foro habilitado en Sicua Plus. Cada grupo debe escoger una competencia diferente, por lo tanto, la elección se dará por el orden de publicación en el foro.

Se tiene libertad sobre el desarrollo del proyecto, para el cual se recomiendan llevar a cabo las siguientes etapas:

Preparación de los datos

Realizar procesos de limpieza que incluyen el tratamiento de Missing values y/o Outliers, manejo de características categóricas.

Evaluación de modelos

La evaluación, calibración y selección de diferentes modelos. Para esta etapa debe tener en cuenta los procesos de Training Score, Validation Score, curva de validación y curva de aprendizaje.

Elección de algoritmos, análisis de resultados y conclusiones**3. Entrega****Presentación Github**

La entrega de resultados se realizará mediante una presentación que deberá ser cargada en la plataforma Github con 3 diapositivas cuyo contenido será:

- Diapositiva 1: Descripción del problema seleccionado.
- Diapositiva 2 y 3: Metodología de abordaje del problema. Donde se muestren los algoritmos utilizados y bajo qué criterios fueron seleccionados.

NOTA: Para la solución del problema es necesario la aplicación de al menos 3 algoritmos.

Notebook Github

Como soporte de la presentación, se deberá cargar a la plataforma Github un Notebook de Jupyter que contenga los algoritmos utilizado para el desarrollo del proyecto.

Bonos

- Bono de hasta 1.0 que dependerá del premio que da o dio la competencia elegida.

- El equipo que satisfactoriamente complete la competencia de mayor premio se le da 1.0 adicional.
A los siguientes dos equipos por valor del premio que satisfactoriamente completen la competencia se les da el bono equivalente a la relación del premio con el mayor premio.