

Viernes 12 de junio de 2020

Datathon2020

Un Datathon es una competencia de tipo "Hackathon" de un fin de semana en la que se reta a los participantes a trabajar en un caso de negocios del mundo real desde diferentes áreas de Machine Learning, Inteligencia Artificial y Ciencia de los Datos.

Los equipos tendrán 48 horas para encontrar una solución a un desafío comercial real basado en los conjuntos de datos proporcionados (Datasets). El jurado y la comunidad otorgarán las soluciones más precisas, pero también las más creativas.

El Datathon se abre a la comunidad global a través de nuestra plataforma. Cualquier persona que quiera unirse a nosotros en cualquiera de nuestros Data Meetups, es bienvenida.

Los Datathoners son parte de la Comunidad de Big Data Dominicana, una comunidad para todos los amantes de los datos, expertos y geeks aprendan unos de otros y apliquen sus conocimientos en la competencia durante el fin de semana.

Durante el Datathon hay 4 roles: participantes (que forman equipos y resuelven los casos), mentores (expertos que ayudan a los equipos a resolver los casos), expertos de la industria (representantes de las empresas que proporcionan los casos con más dominio del tema) y los organizadores.

Las descripciones detalladas de los casos del Datathon se anuncian 1 semana antes del Datathon para que los participantes escojan su desafío, los conjuntos de datos del desafío, se publican el primer día del Datathon (viernes). Los resultados serán compartidos dos días después (domingo).

Los equipos se forman desde el inicio del Datathon. Cada equipo escogerá en qué caso deberá trabajar.

Los equipos pueden usar los datasets que se les proporcionen solo durante la duración del evento. No deben usarlos para fines personales o comerciales fuera del formato y actividades del evento. Con el registro para Datathon, los equipos acuerdan proporcionar también su trabajo (análisis, metodología, código fuente, etc.) a los organizadores para que lo publiquen en la página del evento.

Los equipos presentan su trabajo como un artículo cuando terminan las 48 horas del desafío. Las soluciones de cada equipo se evaluarán con varios criterios: indicadores estadísticos, ingenio de la solución y creatividad de visualización.

El evento termina con los mentores evaluando todas las soluciones de los equipos mediante votación. Para las semifinales, los equipos se asignarán en grupos de 4-5 (trabajando en el mismo caso, y tal vez un grupo con varios casos) el domingo por la tarde. Cada grupo será evaluado por un mentor, un experto de la compañía que proporciona el caso) y un experto de la Comunidad.

El jurado clasificará a todos los equipos dentro del grupo en orden ascendente de cuán buena fue su solución (el rango 1 significa "el mejor"). El equipo que clasifica en el top de su grupo califica para la final.

Los equipos finalistas tendrán una presentación en vivo de 5 minutos seguido de una sesión de preguntas y respuestas de 2 minutos por el jurado. El jurado estará compuesto por mentores y expertos de la Comunidad, los cuales llegarán a clasificar a los finalistas.

El equipo con el mejor rango gana.

Reto 1:

Cuáles factores afectan la propagación de una pandemia? Cuál ha sido la efectividad real de las medidas tomadas? En este reto se utilizarán datos y herramientas de libre acceso para encontrar patrones, visualizar tendencias y explicar los hechos actuales, desde un punto de vista global. Se evaluarán los insights encontrados y las visualizaciones presentadas.

* No se requieren conocimientos de programación!

Reto 2:

El objetivo en este reto es ir más allá, y combinar técnicas de Machine Learning y Epidemiología para predecir la expansión de la pandemia. El reto consiste en utilizar datos de infección, combinarlos con datos socio-demográficos, económicos y ambientales de países durante un mes, para la predicción del ratio de infección promedio en el siguiente mes. Se evaluarán las predicciones utilizando el Median Absolute Error, al igual que las técnicas utilizadas. Se busca que el modelo resultante se pueda utilizar como base para explicar y predecir futuras pandemias, o incluso, su fecha de extinción en un territorio.

* Se recomienda tener nociones de programación para participar en este reto.

Horarios programados para la competencia:

(Creación de los equipos y asignación de trabajos)
Martes 23 de junio – 5:00pm

(Acto de apertura Datathon)
Viernes 26 de junio – 5:00pm

(Desarrollo de la solución con los Datasets)
Sábado 27 de junio – 8:30am – 10:00pm

(Presentaciones finales, selección de los 2 mejores lugares reto. 1 y 2)
Domingo 28 de junio – 5:00pm

Más información acerca de nuestra Comunidad en las siguientes redes sociales:

www.bigdatado.org

<https://www.facebook.com/BigDataDO/>

<https://www.instagram.com/bigdatado/>

<https://twitter.com/BigDataDOM>