

HACKATONS DE BIG DATA

Desarrolla la ciencia de los datos, la programación y las habilidades relacionadas en África en el contexto de enormes proyectos de radioastronomía.

Requisitos y Directrices del Hackathon Astro4Dev

PÁGINA EN DESARROLLO. SI TIENES SUGERENCIAS ESCRIBA A INFO@ASTRO4DEV.ORG

A diferencia de un hackathon convencional, que reúne a profesionales avanzados de la programación y la ciencia de los datos para innovar y resolver problemas complejos, el objetivo del Hackathon Astro4Dev es dar a conocer la ciencia de los datos y el aprendizaje automático, así como algunas de sus técnicas. Los participantes en el hackathon trabajan en interesantes problemas del mundo real basados en datos. Los hackathons están dirigidos a estudiantes de postgrado y jóvenes profesionales de los campos de la ciencia, la tecnología, la ingeniería y las matemáticas (STEM) que tengan un gran interés en la ciencia de los datos y que se beneficien de una mayor exposición y experiencia práctica. El hackathon puede planificarse como un evento presencial o virtual.

Las siguientes indicaciones y requisitos son relevantes para los hackathons de desarrollo de habilidades, como los que se organizan en el marco del **proyecto DARA Big Data**, con el objetivo de proporcionar una exposición a la ciencia de los datos y a las técnicas de aprendizaje automático. Otros hackathons pueden tener formatos diferentes; estas son simplemente directrices que pueden adaptarse en consecuencia.

Requisitos:

- Determine si el evento será presencial, virtual o híbrido.
- En el caso de los eventos presenciales/virtuales, un lugar (por ejemplo, un laboratorio de informática) con suficientes ordenadores de sobremesa/portátiles para que haya al menos un ordenador por equipo.
- En el caso de los proyectos de hackathon que implementan el aprendizaje automático/aprendizaje profundo, debe identificarse una plataforma informática disponible, como un servidor o una nube, a la que los participantes puedan acceder durante el evento, ya que la mayoría de los proyectos de este tipo son intensivos desde el punto de vista computacional. Los participantes también podrían hacer uso de servicios gratuitos de computación en la nube como Google Colab o Kaggle.
- En el caso de los hackathons virtuales, o de los hackathons presenciales en los que el código se ejecutará en una plataforma de computación en la nube, los participantes deberán tener acceso a una conexión a Internet fiable con velocidades de al menos 4 Mbps.
- Un número suficiente de tutores, de manera que haya al menos un tutor por cada dos equipos (cada equipo compuesto por 4-5 participantes). Los tutores deben cumplir los requisitos establecidos en el documento de directrices para tutores.
- Una selección de uno o más proyectos del hackathon de DARA Big Data (ver sección de proyectos más abajo). Los proyectos disponibles abarcan una gama de temas comerciales, científicos y relacionados con el desarrollo.

Lineamientos:

PARTICIPANTES:

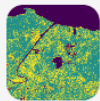
- Determinar las personas que se beneficiarían del evento previsto, es decir, si la escuela/taller/hackathon se dirige a estudiantes (en caso afirmativo, a qué nivel), a educadores, a personas que trabajan y que se beneficiarían de la actualización de sus conocimientos, o una combinación de ambos. En el caso de los eventos virtuales, los participantes necesitan sus propios ordenadores y una conexión fiable a Internet.
- Los participantes deben tener algún nivel de experiencia previa en programación (Python).

FORMAT

- Los hackathones pueden durar un mínimo de dos días (es preferible que sean tres), aunque también pueden celebrarse como parte de un evento más amplio, como un taller o una escuela, en el que se impartan lecciones y/o charlas relacionadas con la ciencia de los datos y el aprendizaje automático. Los hackathones son intensos por naturaleza, en el sentido de que los participantes abarcan mucho contenido y resolución de problemas en un corto espacio de tiempo. Por lo tanto, es importante seguir un programa bien estructurado.
- En el hackathon, los participantes trabajarían con los tutoriales de DARA Big Data antes de abordar la "tarea del hackathon", con la ayuda de tutores que están familiarizados con el material y los conceptos.

HACKATHON PROJECTS

The tutorials relevant to each of the projects below are in the form of [Jupyter Notebooks](#) and, in addition to Python code, contain information, explanations and links to ensure that they are more accessible to the participants. The tutorials also introduce the techniques necessary to complete the hackathon task. Click on any of the following projects to access the GitHub repository with the relevant resources.



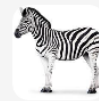
1

Flood Detection



2

NLP - Sentiment Analysis



3

Image Classification



4

Pulsars vs RFI



5

Movie Recommendation App



6

Music Classification

PROYECTOS DE HACKATON

Los tutoriales correspondientes para cada uno de los proyectos que figuran a continuación, están en forma de [cuadernos Jupyter](#) y, además del código Python, contienen información, explicaciones y enlaces para que sean más accesibles a los participantes. Los tutoriales también introducen las técnicas necesarias para completar la tarea del hackathon. Haz clic en cualquiera de los siguientes proyectos para acceder al repositorio de GitHub con los recursos correspondientes.

1. Detecciones de inundaciones.
2. NLP-Análisis del sentimiento
3. Clasificación de imágenes.
4. Pulsares vs RFI
5. Aplicación para recomendar películas.
6. Clasificación de música.

ENTRENAMIENTO DE TUTORES

Para llevar a cabo un hackathon de big data con éxito, se necesitan tutores con experiencia para ayudar a los participantes con los tutoriales del hackathon y la tarea del hackathon. Se recomienda una proporción de un tutor por cada dos equipos del hackathon (cada equipo consta de 4-5 participantes) para garantizar que los tutores no estén sobrecargados y que los participantes reciban asistencia a tiempo. Como preparación para el hackathon, los tutores deben repasar estas [instrucciones para tutores](#) y hacer uso de uno de los vídeos de formación que aparecen a continuación y que son relevantes para el proyecto del hackathon elegido.

TAREA DE HACKATON

- Durante el hackathon, los participantes trabajan en equipos de 4-5 personas para permitir el aprendizaje entre pares y el desarrollo de habilidades blandas como el trabajo en equipo, la comunicación y la colaboración. Se recomienda que cada equipo sea diverso en cuanto a nivel de competencias, formación académica/profesión laboral y género.
- Los participantes podrán presentar sus soluciones hacia el final del evento, mostrando su comprensión de las técnicas utilizadas o demostrando la precisión de sus soluciones. En la medida de lo posible, también se podrán conceder certificados y premios.

INFORMÁTICA

Si el proyecto del hackathon elegido indica que se recomienda una potencia de cálculo adicional para ese proyecto, los participantes pueden hacer uso de la plataforma informática [Google Colab](#).

Los organizadores también pueden ponerse en contacto con el Instituto Interuniversitario de Astronomía con Uso Intensivo de Datos ([IDIA](#)) para obtener un posible apoyo informático mediante el uso de sus instalaciones de investigación en la nube.

RECURSOS

Los **recursos** que pueden ayudar a los participantes en su preparación antes del hackathon, pueden ser distribuidos a todos los participantes aproximadamente dos semanas antes del evento. Estos recursos cubren la programación en Python, así como una introducción básica a ciertos conceptos de aprendizaje automático.

También están disponibles las **charlas grabadas** por miembros de DARA Big Data e IDIA sobre temas de ciencia de datos y aprendizaje automático, que podrán ser reproducidas como parte del evento para beneficio de los participantes. También se recomienda que las charlas invitadas sean impartidas por expertos académicos o de la industria en la materia, con el fin de que los participantes conozcan las múltiples aplicaciones de la ciencia de datos y el aprendizaje automático.

PUBLICIDAD DE EVENTOS

- En el caso de que el evento no se dirija a un grupo preseleccionado de personas, por ejemplo, todos los estudiantes de licenciatura y de máster del departamento de Física de la Universidad, se puede crear un cartel/anuncio (véase el ejemplo de cartel) que contenga toda la información pertinente y que se difunda a través de las listas de correo y las redes sociales con el fin de atraer a los candidatos adecuados.
- Hemos comprobado que incluir un enlace a un formulario de Google en el que las personas pueden presentar su solicitud. Es un método eficaz para obtener inscripciones para el evento.
- Si se requiere un proceso de selección, los participantes pueden ser seleccionados en función de su experiencia en Python o de su motivación para asistir.

COMENTARIOS Y CONEXIONES DESPUÉS DEL EVENTO

- El éxito o el impacto del evento puede evaluarse mediante los comentarios de los participantes y los tutores, y/o la calidad de las soluciones producidas. Aquí se pueden encontrar ejemplos de encuestas previas y posteriores al evento.
- Los organizadores pueden crear redes para que tanto ellos como los participantes, se mantengan en contacto para evaluar el impacto a largo plazo del evento o para llevar a cabo evaluaciones de los participantes después del evento. Estas redes pueden adoptar la forma de listas de correo, grupos de Facebook, grupos de WhatsApp, etc.