

Fundamentos del Aprendizaje Automático para las Ciencias de la Tierra

Parte 2: Ejemplo de Datos de Entrenamiento y Clasificación de la Cobertura Terrestre

Instructores: Jordan A. Caraballo-Vega, Mark L. Carroll, Jules Kouatchou, Jian Li, Caleb S. Spradlin

27 de abril de 2023



Objetivos de esta Capacitación

Al final de la capacitación, los participantes podrán

- Reconocer los métodos de aprendizaje automático más comúnmente utilizados para procesar datos de observación de la Tierra
- Describir los beneficios y las limitaciones del aprendizaje automático para el análisis de datos de observación de la Tierra
- Explicar cómo aplicar algoritmos y técnicas de aprendizaje automático básicos de manera significativa a datos de teledetección
- Usar datos de entrenamiento para evaluar las condiciones y soluciones para un estudio de caso determinado
- Completar los procedimientos básicos para interpretar, refinar y evaluar la precisión de los resultados del análisis de aprendizaje automático

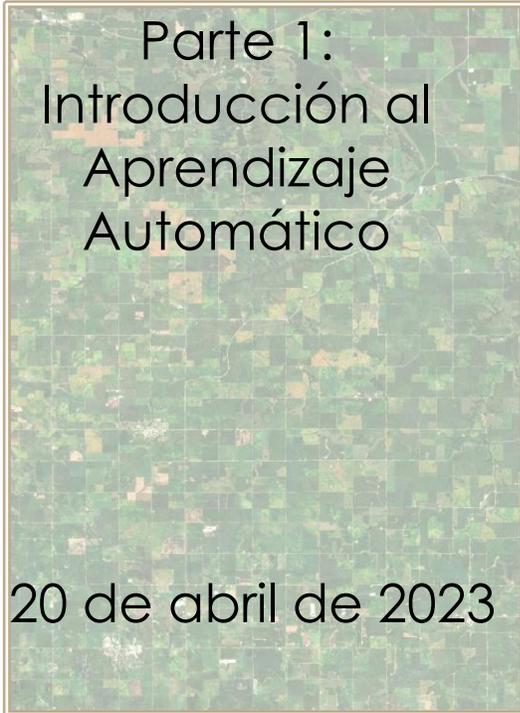


Recordatorio- Pre-requisitos

- Prerrequisitos:
 - La Sesión 1 de nuestra serie disponible a pedido, Fundamentos de la Percepción Remota (Teledetección) o contar con experiencia equivalente (https://appliedsciences.nasa.gov/sites/default/files/2023-02/Fundamentals_of_RS_Span.pdf).
 - Los participantes necesitarán tener acceso a Google Drive y Google Colab. Para acceder a estos recursos, deben utilizar un correo que termine en 'gmail.com'.
 - Pondremos la grabación de esta sesión a su disposición dentro de 48 horas después de la presentación.



Agenda



Parte 1:
Introducción al
Aprendizaje
Automático

20 de abril de 2023

Parte 2:

**Ejemplo de Datos
de Entrenamiento y
Clasificación de la
Cobertura Terrestre**

27 de abril de 2023

Parte 3:

Ajustes de
Modelos,
Optimización de
Parámetros y
Algoritmos de
Aprendizaje
Automático
Adicionales

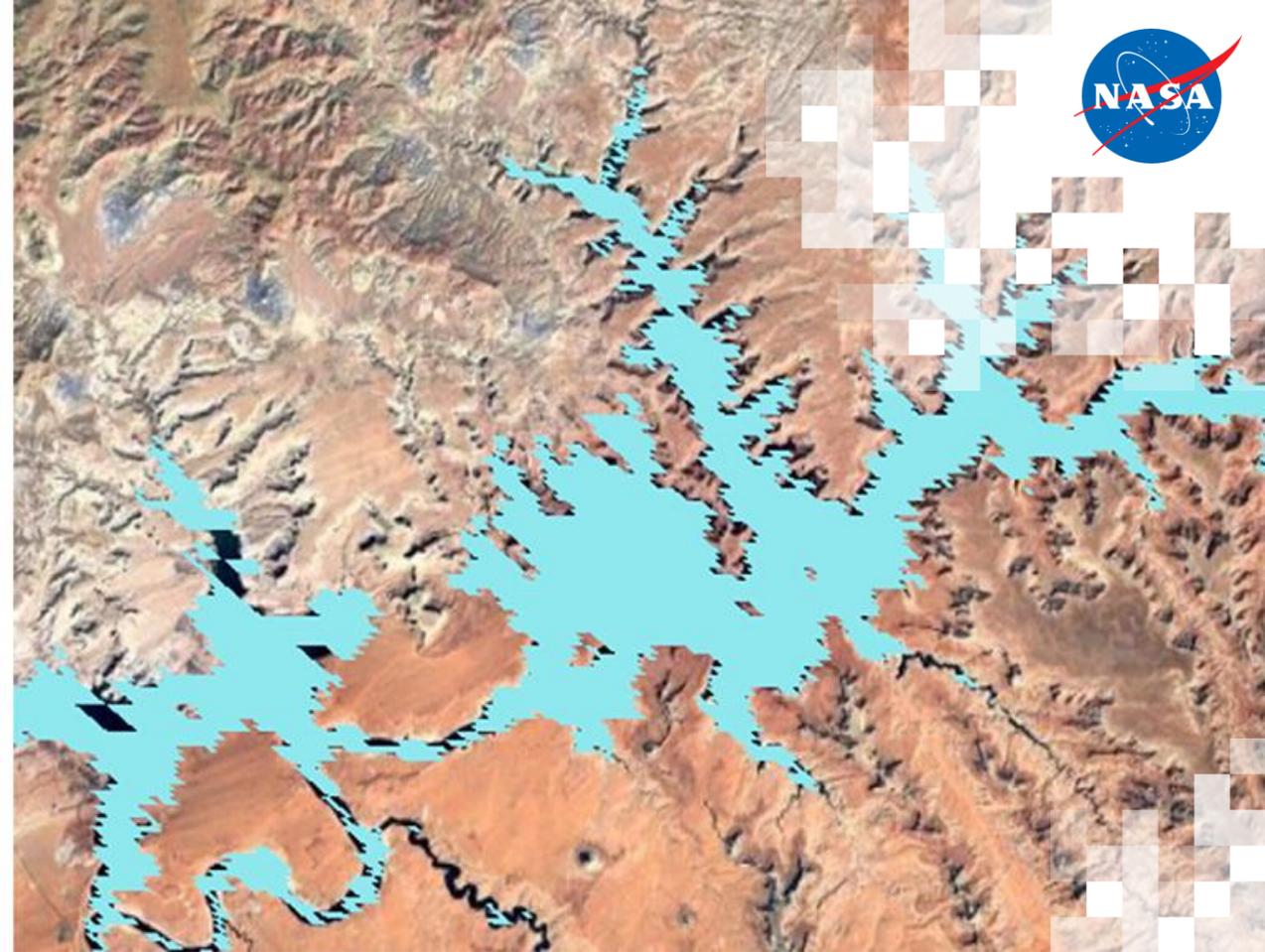
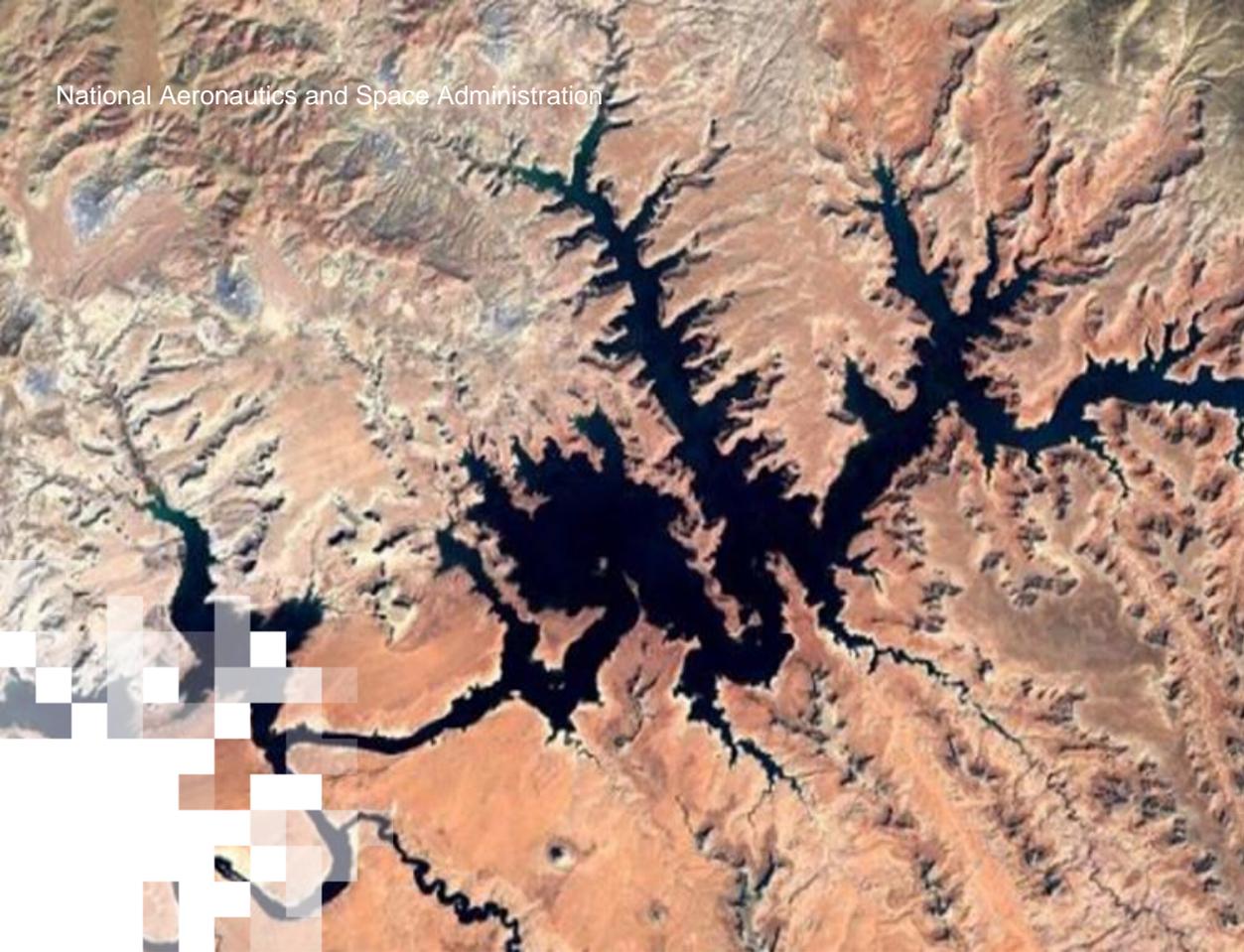
4 de mayo de 2023

Tarea

Práctica y
aplicación
independientes

Fecha límite: 19 de mayo
Disponible: 4 de mayo





Fundamentos del Aprendizaje Automático para las Ciencias de la Tierra

Parte 2: Ejemplo de Datos de Entrenamiento y Clasificación de la Cobertura Terrestre

Instructores: Jordan A. Caraballo-Vega, Mark L. Carroll, Jules Kouatchou, Jian Li, Caleb S. Spradlin

27 de abril de 2023



Esquema de la Sesión 2

- Descargar los datos de entrenamiento
- Análisis de datos exploratorio
- Extracción de datos de entrenamiento de un conjunto de datos tabulares
- Extracción de datos de entrenamiento de un conjunto de datos ráster
- Entrenamiento e inferencia de un conjunto de datos tabulares y ráster
- Métricas y evaluación de modelos
- Ejercicio Práctico en Jupyter Notebook: Estudio de Caso de la Clasificación del Agua de MODIS
- Tarea para después de la Sesión
- Sesión de preguntas y respuestas

Recursos para esta capacitación

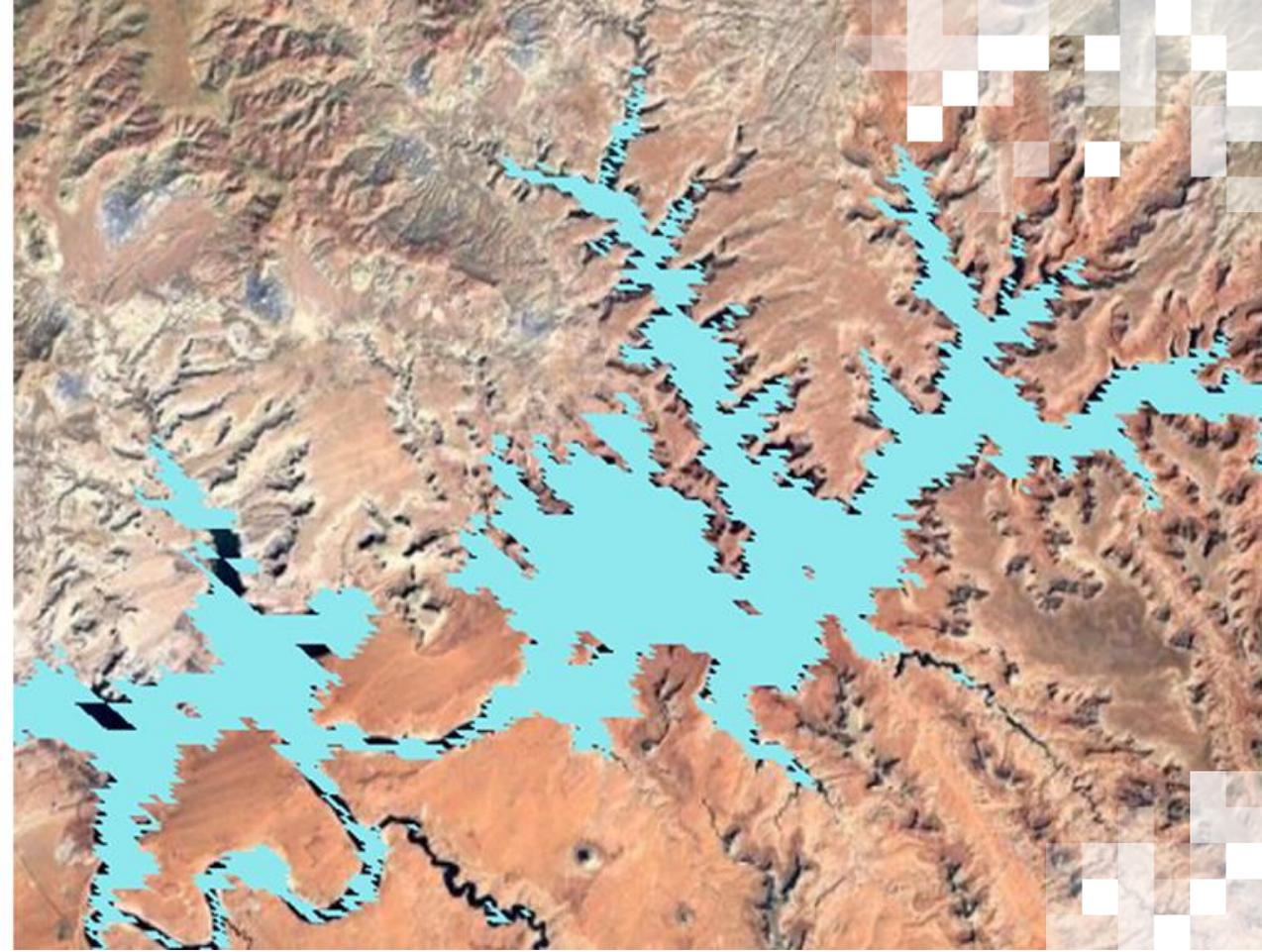
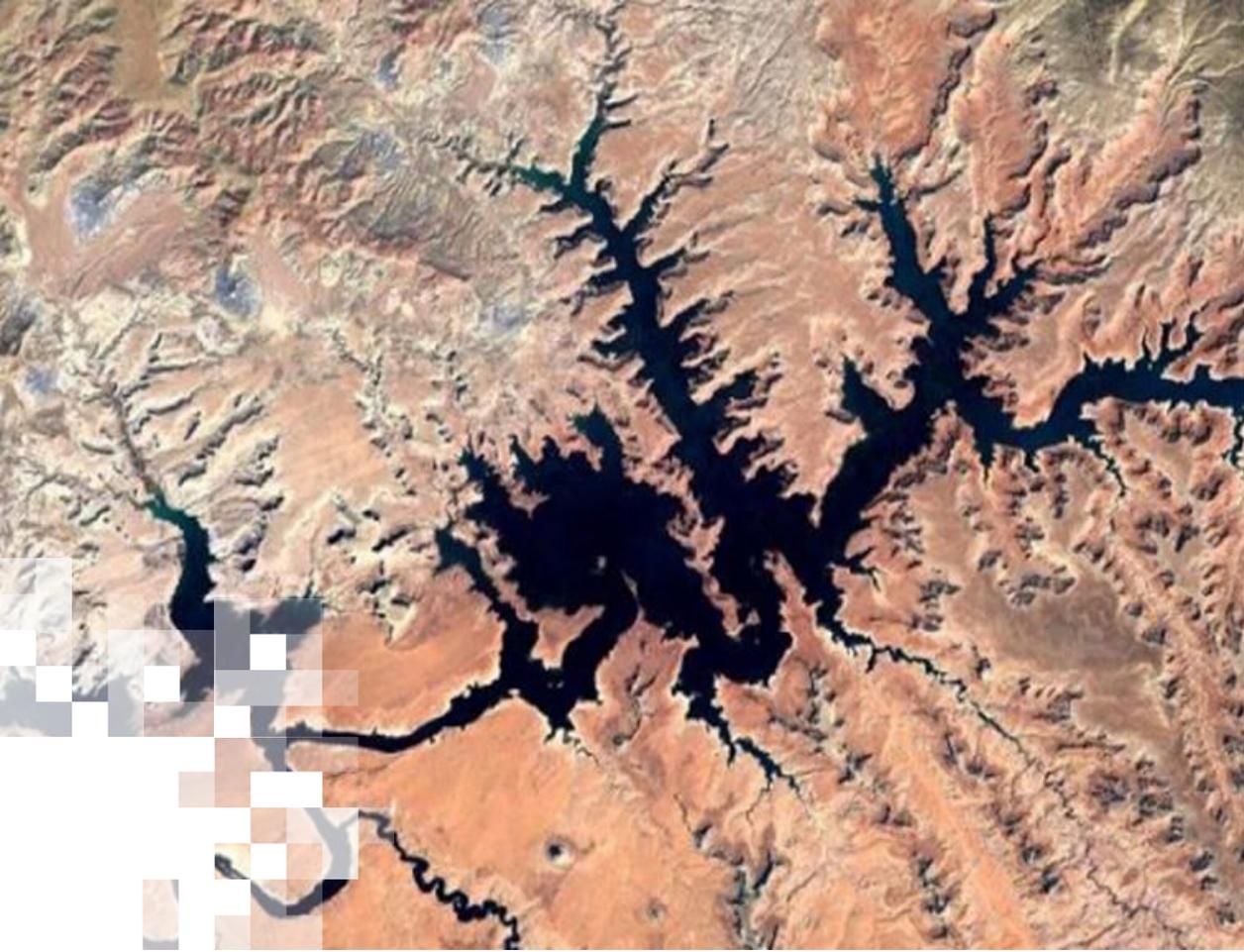
https://github.com/NASAARSET/ARSET_ML_Fundamentals

Objetivos de Capacitación

Después de participar en esta capacitación, quienes asistieron podrán:

- Usar procedimientos de programación básicos para descargar, utilizar y procesar datos de teledetección
- Usar un conjunto de datos de entrenamiento adecuado para el análisis para evaluar las condiciones y soluciones para un estudio de caso determinado
- Completar los procedimientos básicos para interpretar, refinar y evaluar la precisión de los resultados del análisis de aprendizaje automático





Visión General del Instrumento y los Datos

Formador: Jian Li

Moderate Resolution Imaging Spectroradiometer* (MODIS)

- MODIS es un instrumento clave a bordo de los satélites **Terra** y **Aqua**.
- MODIS está avistando la superficie entera de la Tierra cada 1 a 2 días.
- Los productos de datos de MODIS, incluso los de la atmósfera, el océano, las tierras y la criósfera, se utilizan para estudiar los cambios a nivel mundial.
- Los datos adquiridos mejorarán nuestro entendimiento de la dinámica global y los procesos ocurriendo en la tierra, en el océano y en la parte inferior de la atmósfera.

Phytoplankton bloom in the Black Sea in June 2000. Brown sediment discharge from the Danube delta is hugging the western coast, and the phytoplankton bloom is evident by the green and blue colors in the central and eastern side of the image. Image credit: MODIS Land Team/Jacques Desclotres, SSAI; MODIS Ocean Team/Ron Vogel, SAIC/GSC.

NASA's Applied Remote Sensing Training Program

*Espectrorradiómetro de Imágenes de Mediana Resolución, en inglés

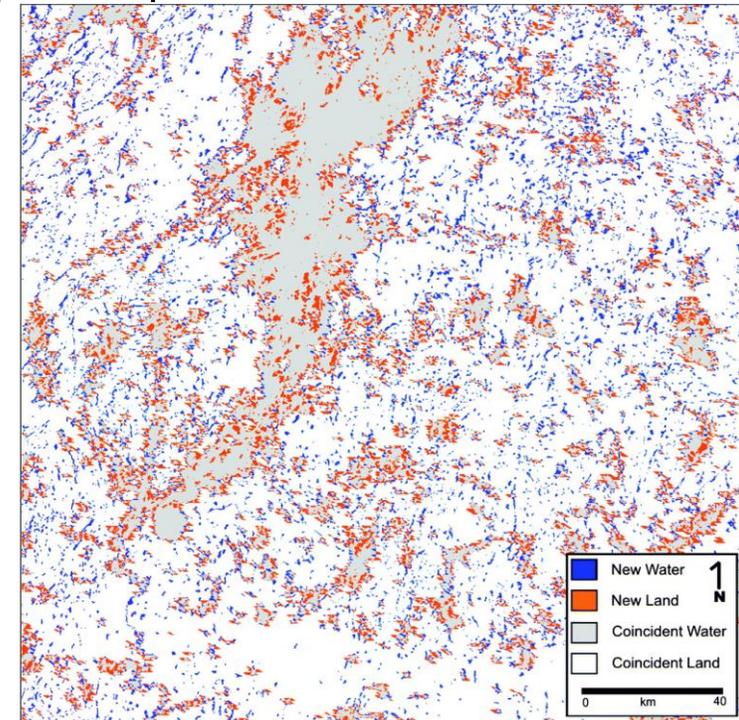
NASA's Earth Observing System



Aplicaciones de Productos de MODIS

Máscara de Agua Global Ráster a Resolución de 250m

- Expandir la cobertura espacial para incluir el planeta entero y abordar algunas discontinuidades erróneas en las redes fluviales principales



Flujo de Trabajo de la Clasificación del Agua de MODIS

Representación mejorada de los lagos en Canadá boreal al oeste de la bahía de Hudson comparada con la máscara de agua de EOS antigua.

Fuente de la Imagen: Carroll et al.

<https://doi.org/10.1080/17538940902951401>



Acceso a Datos de MODIS

- Los datos de la NASA están almacenados en Centros de Archivos Activos Distribuidos (Distributed Active Archive Centers o DAACs).
- Datos MODIS de 1^{er} Nivel, Productos de Geolocalización, Máscara de Nubes y la Atmósfera:
 - <http://ladsweb.nascom.nasa.gov/>
- Productos de Tierras de MODIS:
 - <https://lpdaac.usgs.gov/>
- Productos de la Criósfera de MODIS:
 - <http://nsidc.org/daac/modis/index.html>
- Productos del Color Oceánico y la Temperatura Superficial Marina de MODIS:
 - <http://oceancolor.gsfc.nasa.gov/>



NASA Earthdata Search

Earthdata Search brinda acceso a todos los datos de los DAAC vía una interfaz en un mapa en la web.

Buscar datos según criterios como instrumento

The screenshot shows the NASA Earthdata Search interface. On the left, a sidebar lists various instruments, with 'MODIS' selected. The main panel displays a list of 955 matching collections. Several MODIS products are highlighted, including 'MODIS/Aqua Surface Reflectance Daily L2G Global 1km and 500m SIN Grid V006', 'MODIS/Aqua Surface Reflectance Daily L2G Global 250m SIN Grid V006', 'MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V061', 'MODIS/Terra Surface Reflectance Daily L2G Global 1km and 500m SIN Grid V061', 'MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V061', and 'MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006'. A world map on the right shows the spatial distribution of the data.

Lista de Productos de MODIS

Vista Previa de Datos de Muestreo



NASA Earthdata Search, Continuación

Filtrar datos según el nombre del producto, cobertura espacial, intervalo de tiempo etc.

The screenshot displays the NASA Earthdata Search interface. On the left, a sidebar contains filter options for Spatial, Temporal, and Data Access. The main panel shows search results for 'MODIS/Aqua Surface Reflectance Daily L2G Global 1km and 500m SIN Grid V006'. Two granules are listed with their respective start and end times. A map on the right shows the search area over the United States and Mexico, with a zoomed-in view of a specific data sample.

Lista de Gránulos de MODIS

Vista Previa de Datos de Muestreo



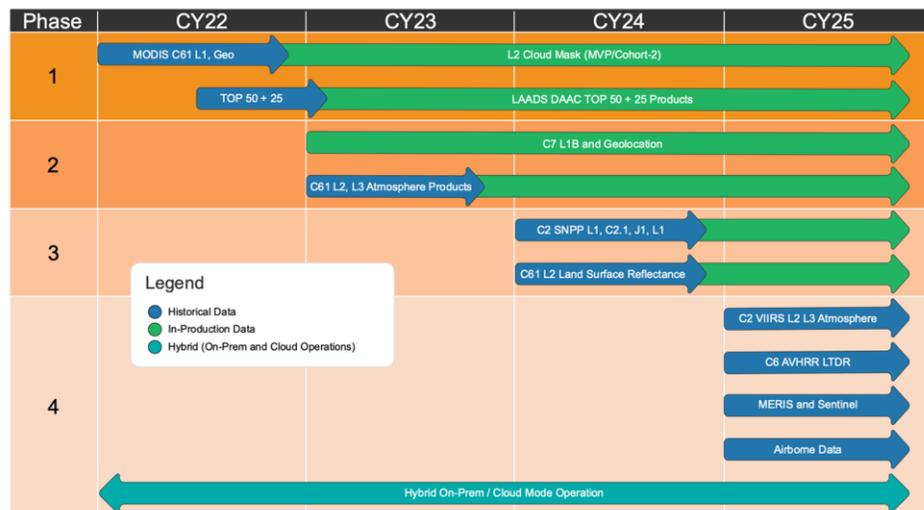
Datos MODIS en la Nube

Como parte de la política de ciencia abierta de la NASA y otras metas relacionadas, todos los DAACs están migrando sus colecciones a [Earthdata Cloud](#).

The Amazon Sustainability Data Initiative
ASDI

MODIS MYD13A1, MOD13A1, MYD11A1, MOD11A1, MCD43A4

This product is part of the Amazon Sustainability Data Initiative and contains data sets that are publicly available for anyone to access and use. No subscription is required. Unless specifically stated in the applicable data set documentation, data sets available through the Amazon Sustainability Data Initiative are not provided and maintained by AWS.



Description Resources on AWS Usage examples Links Similar products

Contact: <https://astraea.earth/>

General AWS Data Exchange support
Contact Us [↗](#)

Resources on AWS

Description

Imagery and metadata

Resource type

S3 Bucket

Amazon Resource Name (ARN)

`arn:aws:s3:::astraea-opendata`

AWS Region

us-west-2

AWS CLI Access (No AWS account required)

`aws s3 ls --request-payer requester s3://astraea-opendata/`

Los cronogramas propuestos para migrar los productos de datos del LAADS DAAC (sup. izq.)

Muestras de productos de MODIS migrados en la fase temprana (inf. izq.)

Fuente: LAADS DAAC

<https://ladsweb.modaps.eosdis.nasa.gov/cloud/>

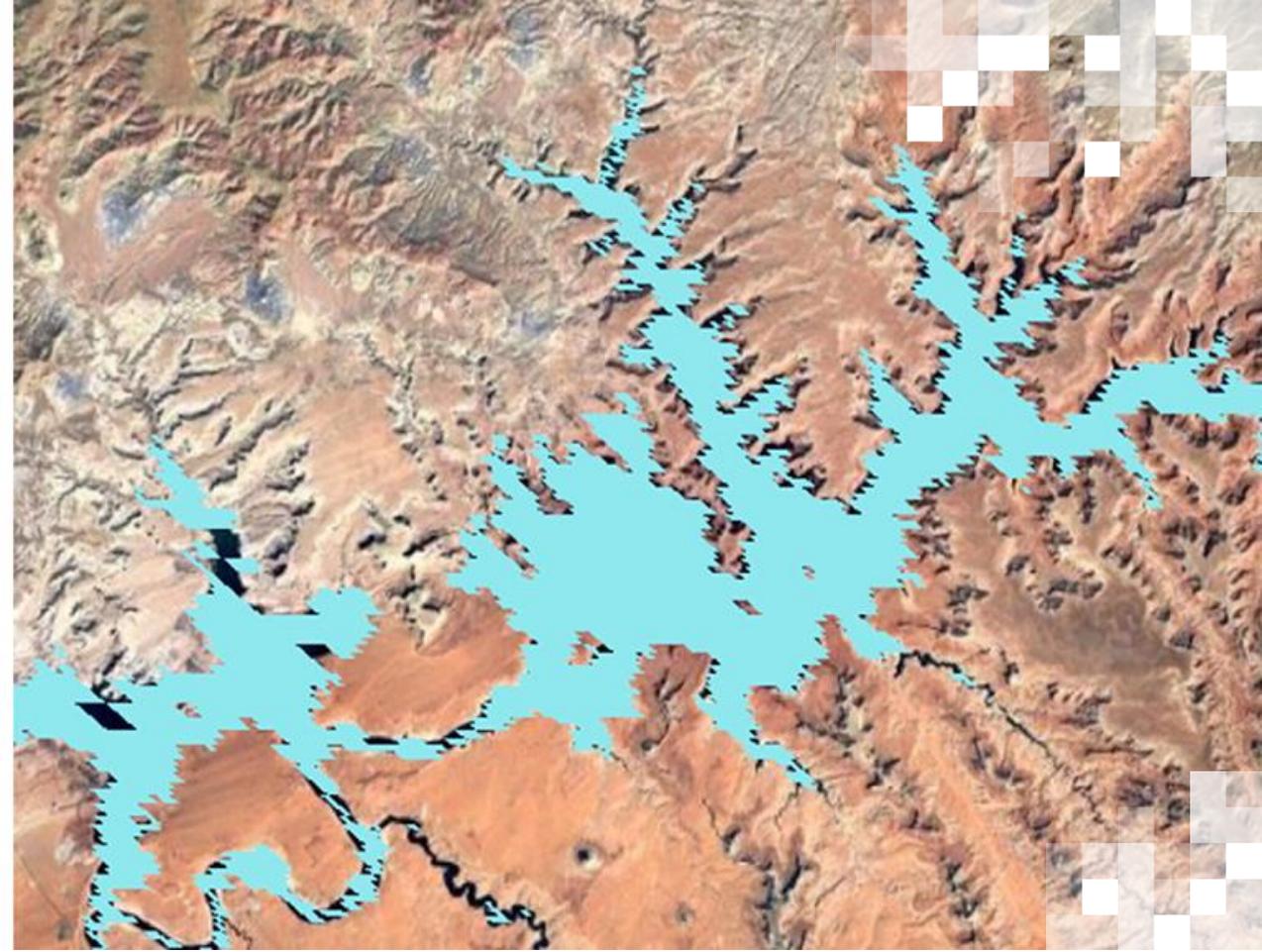
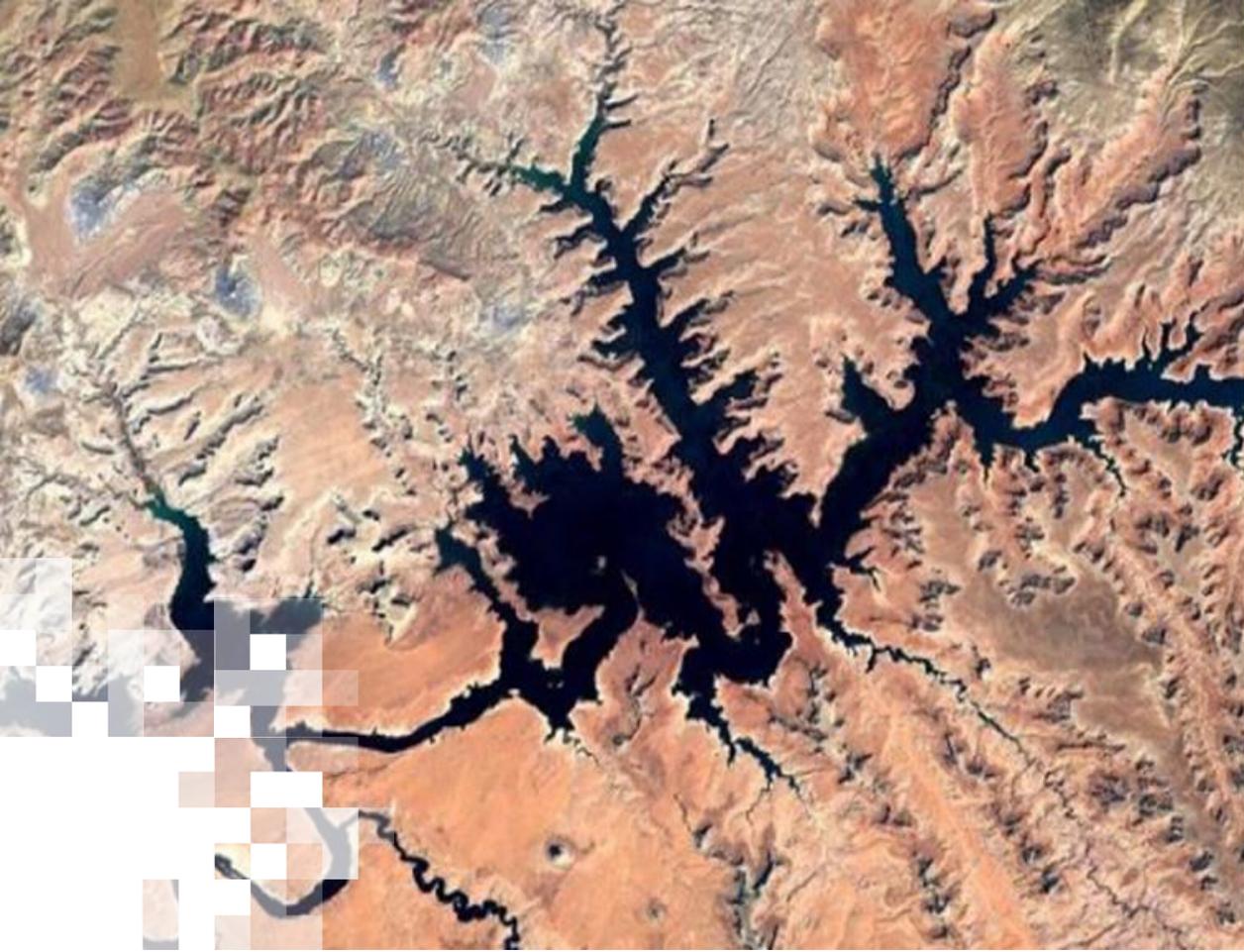
Productos de MODIS disponibles a través de ASDI (der.)

Fuente: AWS ASDI

<https://registry.opendata.aws/modis-astraea/>

LAADS DAAC Phase 1.1 (Cohort-2) Datasets					
Shortname	Platform	Instrument	Description	Availability	
MOD021KM	Terra	MODIS	Level 1B Calibrated Radiances - 1km	December 2022	
MYD021KM	Aqua	MODIS	Level 1B Calibrated Radiances - 1km	December 2022	
MOD02HKM	Terra	MODIS	Level 1B Calibrated Radiances - 500m	December 2022	
MYD02HKM	Aqua	MODIS	Level 1B Calibrated Radiances - 500m	December 2022	
MOD02QKM	Terra	MODIS	Level 1B Calibrated Radiances - 250m	December 2022	
MYD02QKM	Aqua	MODIS	Level 1B Calibrated Radiances - 250m	December 2022	
MOD03	Terra	MODIS	Geolocation - 1km	December 2022	
MYD03	Aqua	MODIS	Geolocation - 1km	December 2022	
MOD35_L2	Terra	MODIS	Cloud Mask and Spectral Test Results 5-Min L2 Swath 250 and 1km	December 2022	





Análisis de Datos Exploratorio

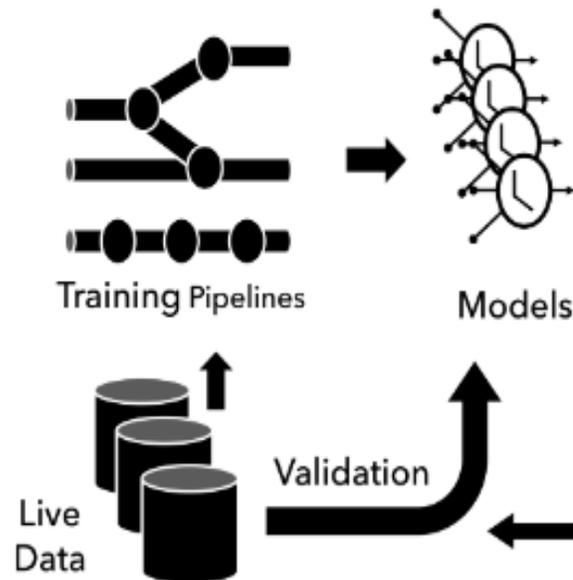
Formador: Caleb S. Spradlin

Canal de Desarrollo del Aprendizaje Automático de la Sesión 1

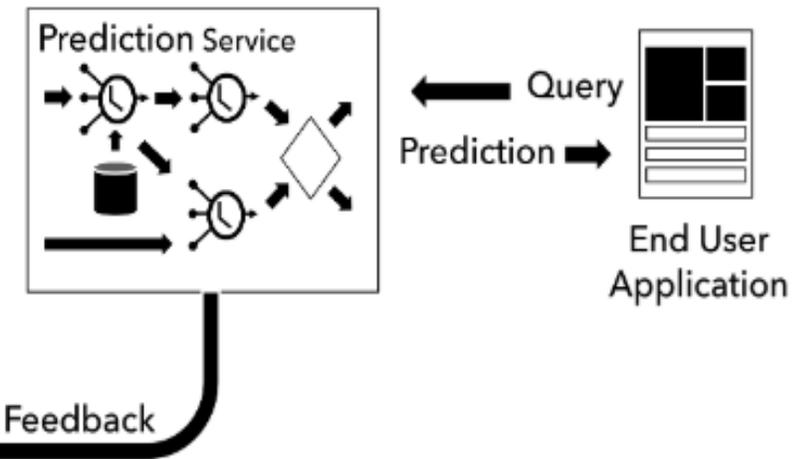
Pipeline Development



Training



Inference

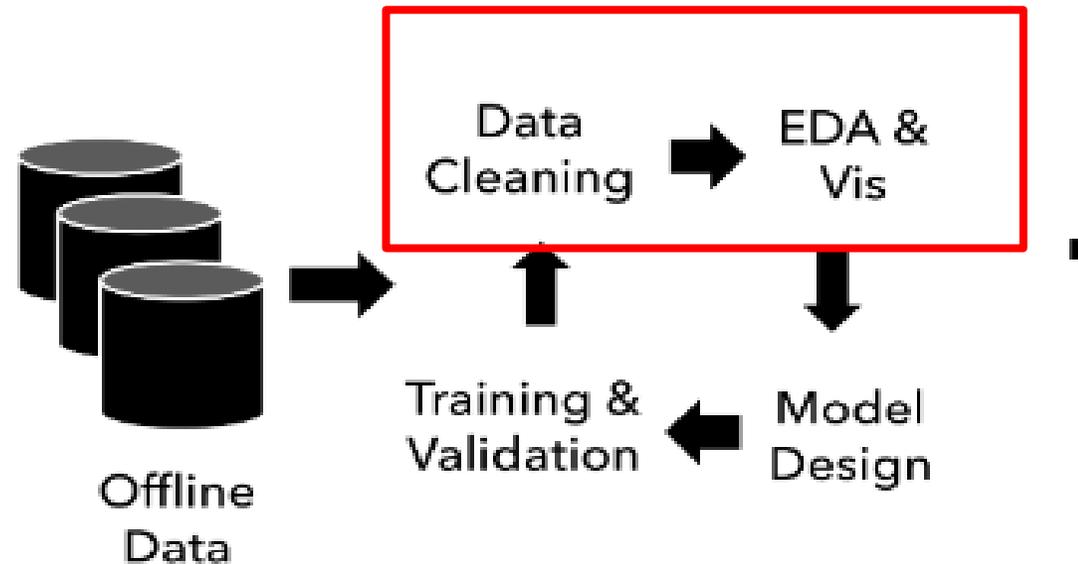


Fuente de la Imagen: Daniel Crankshaw (En: 'A Short History of Prediction-Serving Systems')



Análisis de Datos Exploratorio (Exploratory Data Analysis o EDA) para el Aprendizaje Automático

Pipeline Development



Fuente de la Imagen: Daniel Crankshaw (En: 'A Short History of Prediction-Serving Systems')



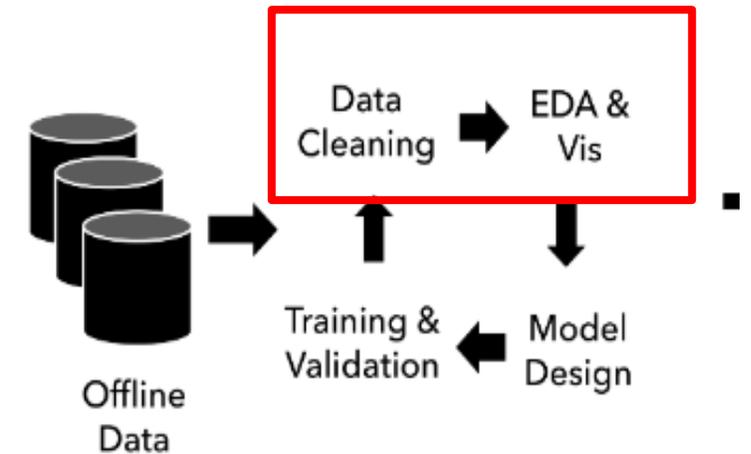
Análisis de Datos Exploratorio (Exploratory Data Analysis o EDA) para el Aprendizaje Automático

El EDA es un enfoque para el análisis de datos que utiliza una variedad de técnicas para obtener información sobre los datos.

Pasos básicos en el análisis de datos exploratorio:

- Limpieza y preprocesamiento
- Análisis estadístico
- Visualización para el análisis de tendencias, detección de anomalías, detección (y eliminación) de valores atípicos

Pipeline Development



Fuente de la Imagen: Daniel Crankshaw (En: 'A Short History of Prediction-Serving Systems')



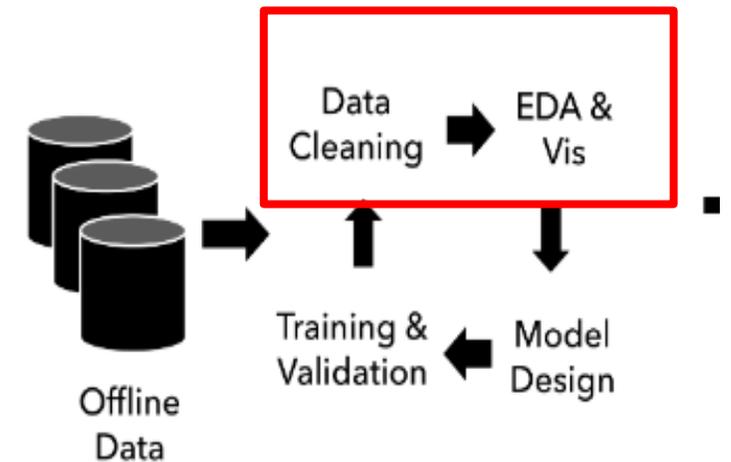
EDA – Importancia del EDA

Mejorar nuestro entendimiento de la estructura y las propiedades del conjunto de datos

Descubrir errores, valores faltantes y valores atípicos en el conjunto de datos

Identificar correlaciones y patrones al visualizar datos

Pipeline Development

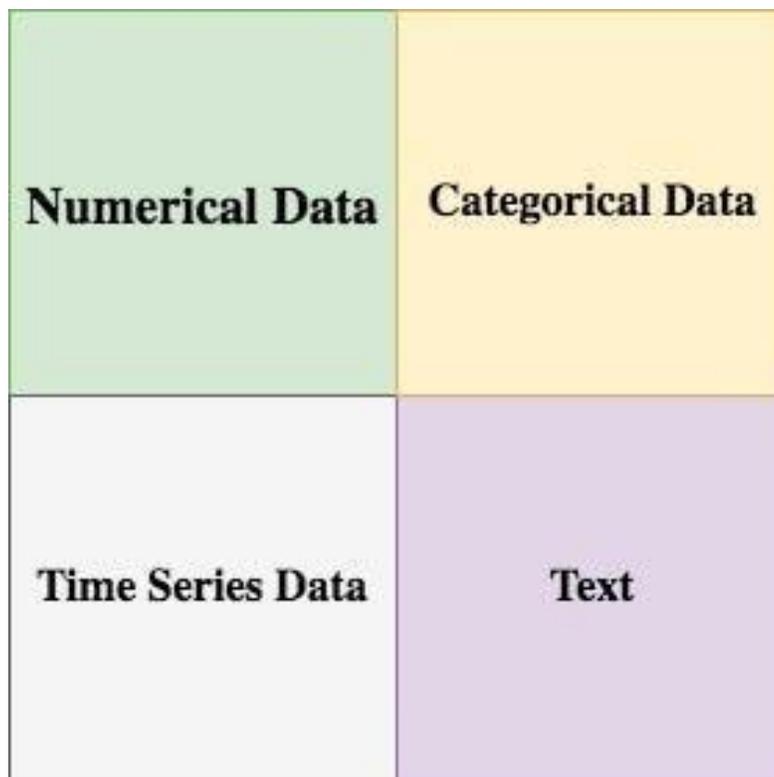


Fuente de la Imagen: Daniel Crankshaw (En: 'A Short History of Prediction-Serving Systems')



EDA – Entendiendo los Diferentes Tipos de Datos

Los datos pueden tener muchas formas.



Tipos de Datos Comunes en el Aprendizaje Automático

```
df.info()

<class 'pandas.core.frame.DataFrame'>
Int64Index: 1007000 entries, 0 to 1006999
Data columns (total 11 columns):
#   Column                Non-Null Count  Dtype
---  ---                -
0   water                 1007000 non-null  Int64
1   sur_refl_b01_1       1007000 non-null  Int64
2   sur_refl_b02_1       1007000 non-null  Int64
3   sur_refl_b03_1       1007000 non-null  Int64
4   sur_refl_b04_1       1007000 non-null  Int64
5   sur_refl_b05_1       1007000 non-null  Int64
6   sur_refl_b06_1       1007000 non-null  Int64
7   sur_refl_b07_1       1007000 non-null  Int64
8   ndvi                  1007000 non-null  Float32
9   ndwi1                 1007000 non-null  Float32
10  ndwi2                 1007000 non-null  Float32
dtypes: Float32(3), Int64(8)
memory usage: 91.2 MB
```

Información de Pandas DataFrame del Conjunto de Datos de Entrenamiento de MODIS



EDA – Limpieza de Datos y Manejo de Datos Faltantes

- Los valores faltantes pueden representar un desafío para los modelos de aprendizaje automático, por lo que es importante comprender la cantidad de datos faltantes en el conjunto de datos.
- Se pueden usar diferentes estrategias, como la imputación o la eliminación, para manejar los valores faltantes según la cantidad y la naturaleza de los datos faltantes.

Detección

Detección de Valores Nulos:

- `isnull()`: Sirve de alias para `dataframe.isna()`. Esta función devuelve el marco de datos con valores booleanos que indican valores faltantes.
- Sintaxis:
`dataframe.isnull()`

Manejo

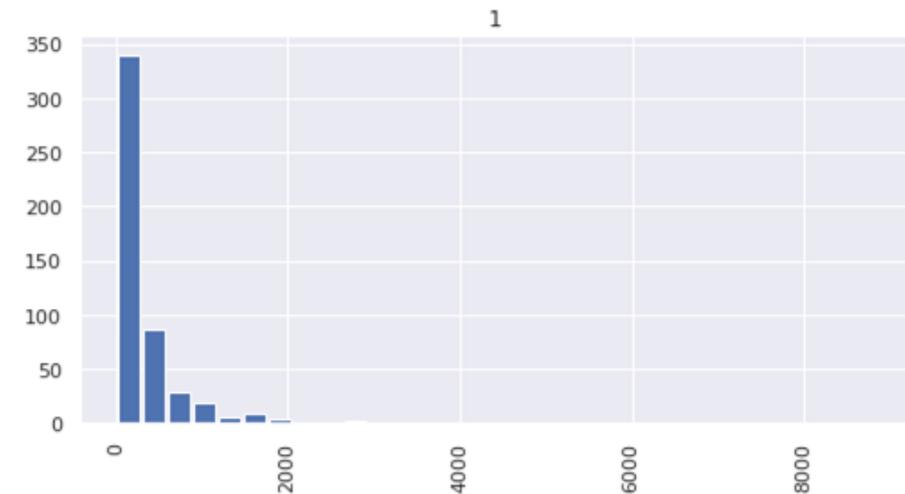
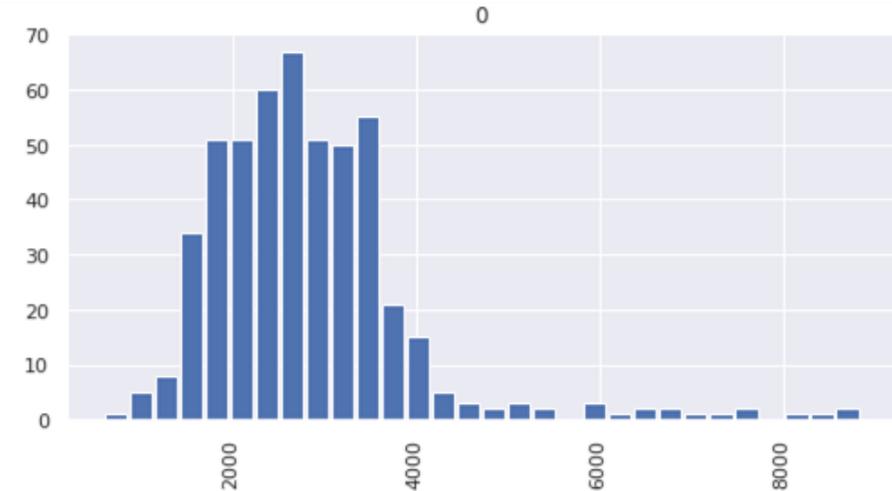
Manejo de Valores Nulos:

- Dar de baja las filas con valores nulos: la función `dropna()` sirve para borrar filas o columnas con valores nulos.
- Reemplazar valores faltantes: la función `fillna()` puede llenar los valores faltantes con un valor especial como una media o un promedio.



EDA – Visualización de Distribuciones de Datos

- Uno de los primeros pasos en el EDA es comprender la distribución de cada variable en el conjunto de datos.
- Los histogramas, las diagramaciones de densidad, las diagramaciones de caja y las diagramaciones de violín se usan comúnmente.
- Comprender las distribuciones de datos puede ayudar a identificar valores atípicos, sesgos y transformaciones posibles que pueden ser necesarias.

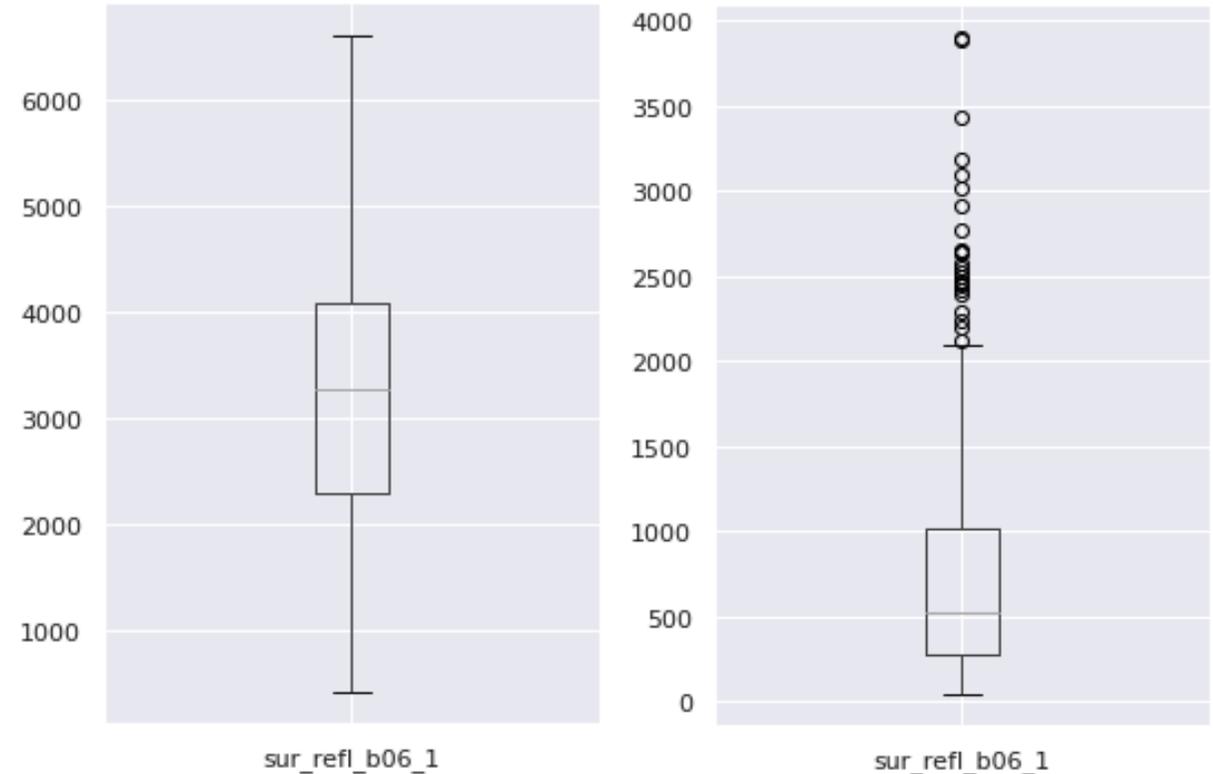


Ocurrencia de Clases en Diferentes Áreas de Respuesta Espectral



EDA – Detección y Tratamiento de Valores Atípicos

- Los valores atípicos pueden tener un impacto sobre el rendimiento de los modelos de ML.
- Las diagramaciones de caja (Boxplots) y las de dispersión (scatterplots) se pueden utilizar para identificar valores atípicos.



Distribución Espectral de Dos Bandas de Reflectancia Superficial de MODIS



EDA – Investigación de Correlaciones

- Las correlaciones entre variables pueden proporcionar información sobre las relaciones y dependencias dentro del conjunto de datos.
- El identificar correlaciones fuertes puede ayudar con la selección de características y puede informar la elección de algoritmos de aprendizaje automático.

	water	sur_refl_b01_1	sur_refl_b02_1	sur_refl_b03_1	sur_refl_b04_1	sur_refl_b05_1	sur_refl_b06_1	sur_refl_b07_1	ndvi	ndwi1	ndwi2
water	1.000000	-0.257704	-0.569499	-0.170032	-0.220100	-0.714663	-0.740303	-0.303113	-0.490893	-0.348729	-0.282386
sur_refl_b01_1	-0.257704	1.000000	0.895279	0.977558	0.989731	0.467660	0.167929	0.101960	0.073173	0.453949	0.292324
sur_refl_b02_1	-0.569499	0.895279	1.000000	0.851850	0.880881	0.701883	0.425545	0.198049	0.299217	0.541591	0.387248
sur_refl_b03_1	-0.170032	0.977558	0.851850	1.000000	0.995246	0.392647	0.040463	0.053476	0.018081	0.396507	0.253936
sur_refl_b04_1	-0.220100	0.989731	0.880881	0.995246	1.000000	0.436087	0.099287	0.076514	0.055002	0.436115	0.283203
sur_refl_b05_1	-0.714663	0.467660	0.701883	0.392647	0.436087	1.000000	0.748920	0.372319	0.421938	0.379261	0.291508
sur_refl_b06_1	-0.740303	0.167929	0.425545	0.040463	0.099287	0.748920	1.000000	0.384666	0.420780	0.211944	0.169933
sur_refl_b07_1	-0.303113	0.101960	0.198049	0.053476	0.076514	0.372319	0.384666	1.000000	0.195150	0.126411	0.149680
ndvi	-0.490893	0.073173	0.299217	0.018081	0.055002	0.421938	0.420780	0.195150	1.000000	0.391189	0.317795
ndwi1	-0.348729	0.453949	0.541591	0.396507	0.436115	0.379261	0.211944	0.126411	0.391189	1.000000	0.485096
ndwi2	-0.282386	0.292324	0.387248	0.253936	0.283203	0.291508	0.169933	0.149680	0.317795	0.485096	1.000000

Coeficientes de Correlación para Diferentes Bandas de Reflectancia Superficial de MODIS



EDA – Investigación de Correlaciones

	water	sur_refl_b01_1	sur_refl_b02_1	sur_refl_b03_1	sur_refl_b04_1	sur_refl_b05_1	sur_refl_b06_1	sur_refl_b07_1	ndvi	ndwi1	ndwi2
water	1.000000	-0.257704	-0.569499	-0.170032	-0.220100	-0.714663	-0.740303	-0.303113	-0.490893	-0.348729	-0.282386
sur_refl_b01_1	-0.257704	1.000000	0.895279	0.977558	0.989731	0.467660	0.167929	0.101960	0.073173	0.453949	0.292324
sur_refl_b02_1	-0.569499	0.895279	1.000000	0.851850	0.880881	0.701883	0.425545	0.198049	0.299217	0.541591	0.387248
sur_refl_b03_1	-0.170032	0.977558	0.851850	1.000000	0.995246	0.392647	0.040463	0.053476	0.018081	0.396507	0.253936
sur_refl_b04_1	-0.220100	0.989731	0.880881	0.995246	1.000000	0.436087	0.099287	0.076514	0.055002	0.436115	0.283203
sur_refl_b05_1	-0.714663	0.467660	0.701883	0.392647	0.436087	1.000000	0.748920	0.372319	0.421938	0.379261	0.291508
sur_refl_b06_1	-0.740303	0.167929	0.425545	0.040463	0.099287	0.748920	1.000000	0.384666	0.420780	0.211944	0.169933
sur_refl_b07_1	-0.303113	0.101960	0.198049	0.053476	0.076514	0.372319	0.384666	1.000000	0.195150	0.126411	0.149680
ndvi	-0.490893	0.073173	0.299217	0.018081	0.055002	0.421938	0.420780	0.195150	1.000000	0.391189	0.317795
ndwi1	-0.348729	0.453949	0.541591	0.396507	0.436115	0.379261	0.211944	0.126411	0.391189	1.000000	0.485096
ndwi2	-0.282386	0.292324	0.387248	0.253936	0.283203	0.291508	0.169933	0.149680	0.317795	0.485096	1.000000

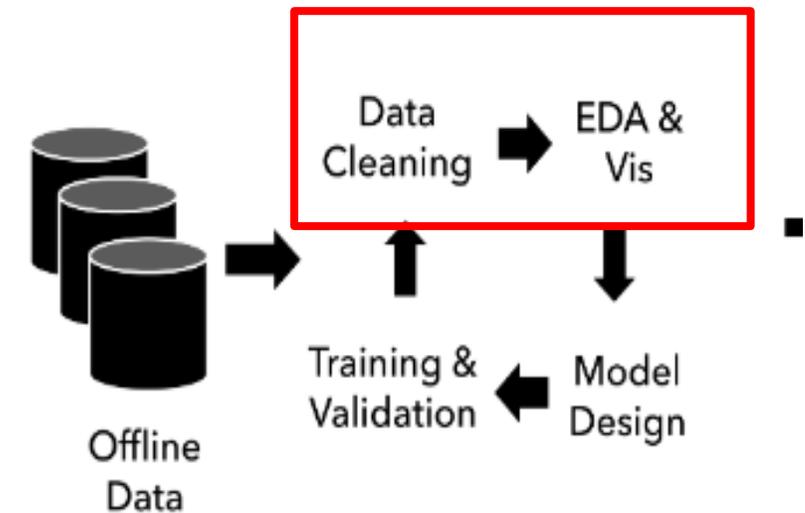
Coeficientes de Correlación para Diferentes Bandas de Reflectancia Superficial de MODIS



EDA para el Aprendizaje Automático – Conclusión

- El EDA es un paso crítico en cualquier proyecto de aprendizaje automático para comprender la estructura y las propiedades de los conjuntos de datos.
- Las visualizaciones y las técnicas de análisis, como las distribuciones de datos, las correlaciones, el manejo de valores faltantes y la detección de valores atípicos, pueden proporcionar información valiosa para la selección de características, el preprocesamiento y la selección de modelos.
- Con estas técnicas y herramientas de EDA, se pueden desarrollar modelos de aprendizaje automático con una mejor comprensión de los datos subyacentes.

Pipeline Development



Extracción de Datos de Entrenamiento de un Conjunto de Datos Tabulares

- Ahora entendemos y hemos limpiado nuestros datos. ¿Qué sigue?

Mejorar nuestro entendimiento de la estructura y las propiedades del conjunto de datos

Descubrir errores, valores faltantes y valores atípicos en el conjunto de datos

Identificar correlaciones y patrones al visualizar datos



Datos de Muestreo

A veces es un reto recopilar suficientes datos para que un modelo de ML generalice.

Sin embargo, a veces, este no es el problema en la teledetección.

Las grandes cantidades de datos traen preguntas.

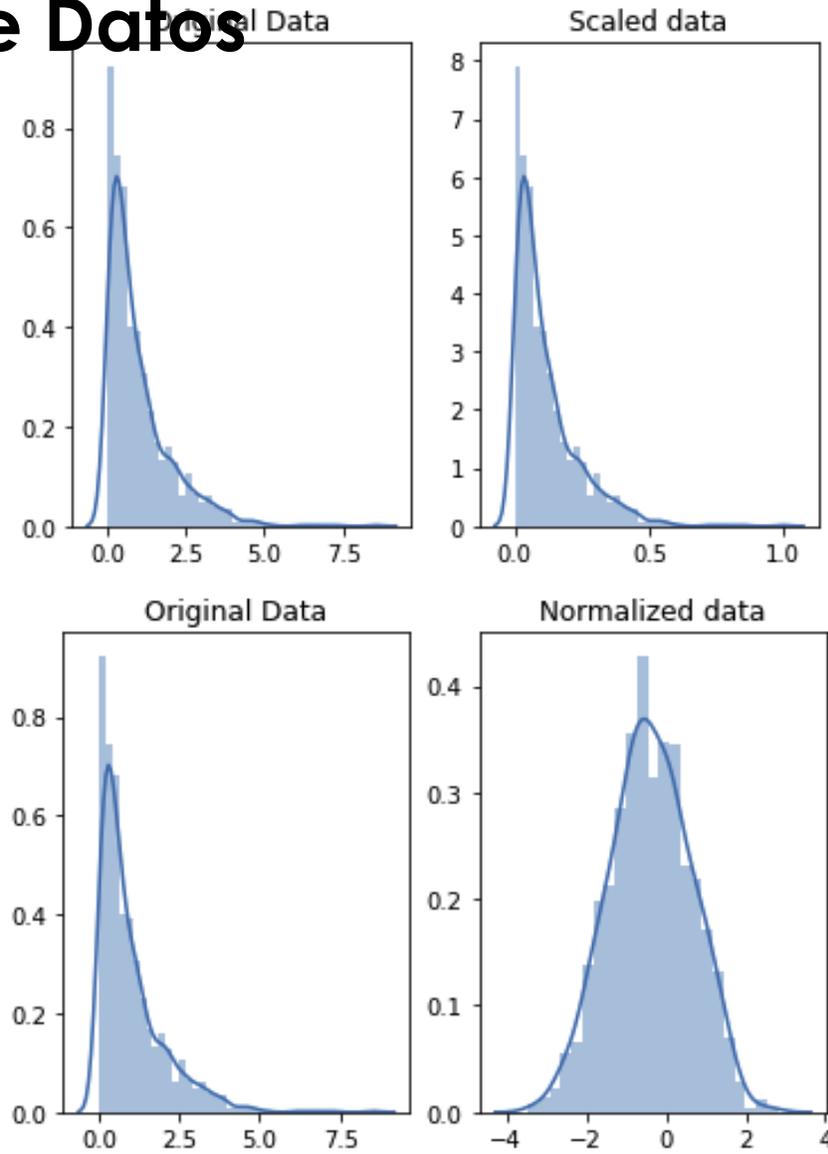
Preguntas:

- ¿Son demasiados datos para un modelo de ML?
- ¿Qué tipo de muestreo realizamos?
- ¿A qué granularidad muestreamos?



Regularización y Escalamiento de Datos

- Escalar y regularizar sus datos puede ayudar a mejorar el rendimiento del modelo.
- El escalamiento y la regularización se deben realizar en los datos a los que se aplica el modelo, no solo en los datos de entrenamiento. Recuerde la regla de oro: **los datos de prueba deben coincidir lo más posible con los datos de entrenamiento.**

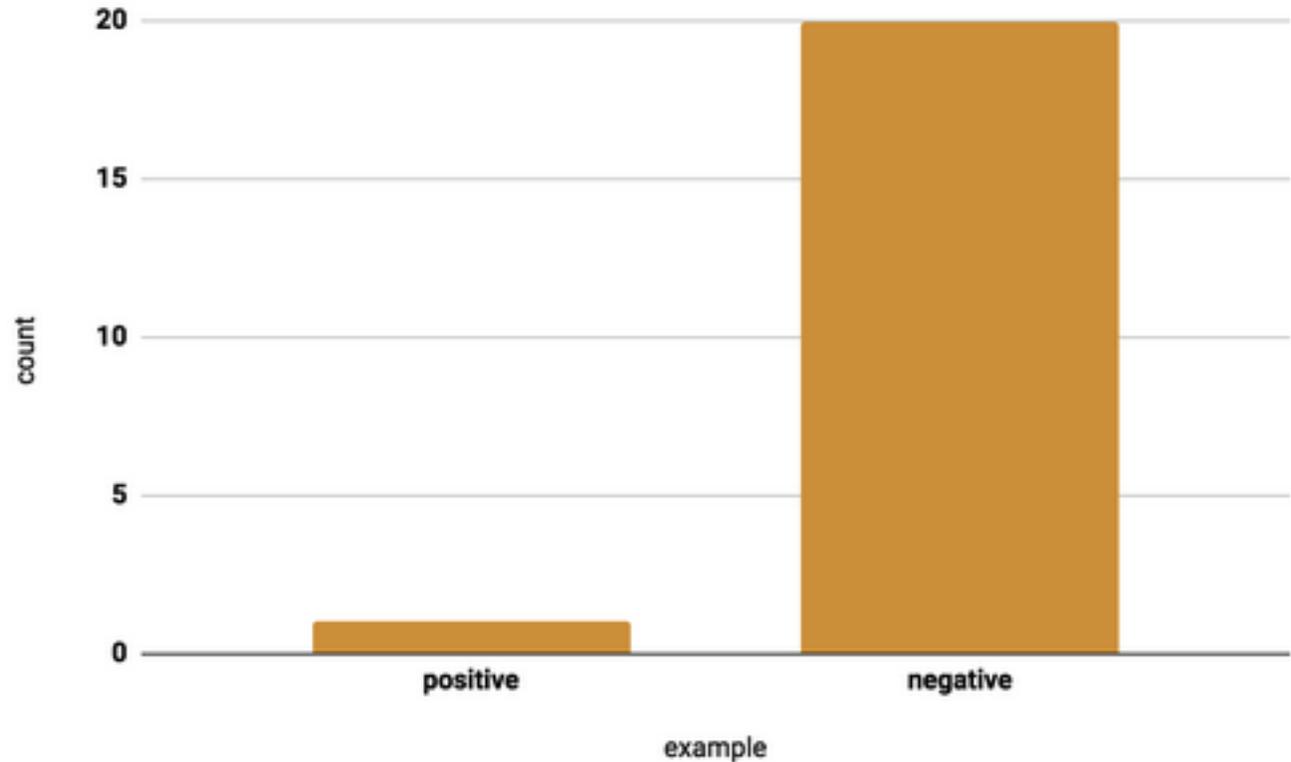


Comparación de Datos Originales y Datos Reescalados y Normalizados



Cómo Manejar Datos Desbalanceados

- Maneras efectivas de manejar datos desbalanceados:
 - Downsampling (reducción de resolución de muestreo)
 - Upweighting (ponderación)
- Algoritmos de ML que pueden manejar pesos en las muestras:
 - Decision Trees
 - Random Forest
 - Gradient Boosting
 - Neural Network



Fuente de la Imagen:
developers.google.com



Partición de Datos

Parta sus datos en dos subconjuntos:

- Conjunto de entrenamiento: Un subconjunto para entrenar un modelo
- Conjunto de validación: Un subconjunto para evaluar el rendimiento durante el entrenamiento
- Conjunto de prueba: Un subconjunto para probar el modelo entrenado

💡 To design a split that is representative of your data, consider what the data represents. The golden rule applies to data splits as well: the testing task should match the production task as closely as possible.

Relación General de Conjuntos de Datos de Entrenamiento, Validación y Prueba



Conjunto de Entrenamiento

Conjunto de Validación

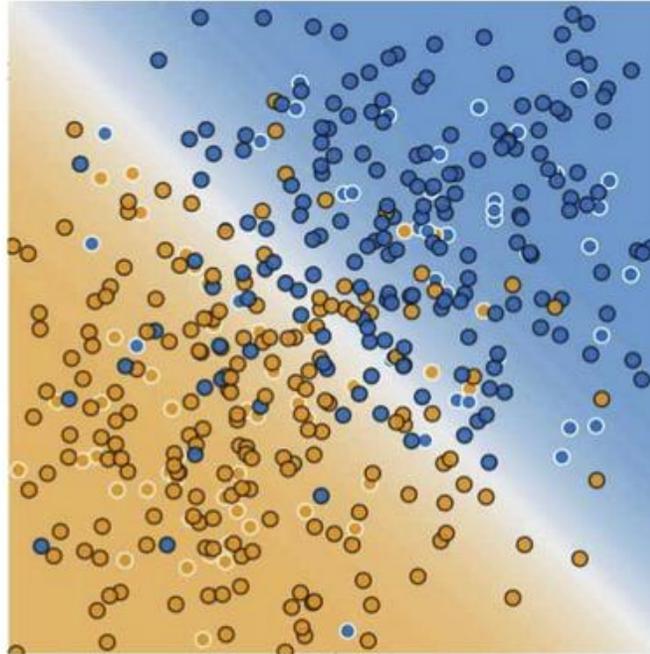
Conjunto de Prueba



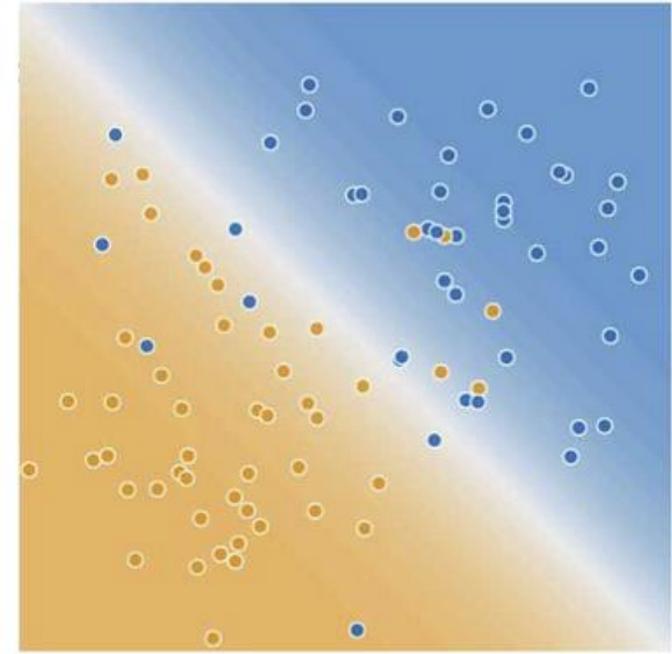
Partición de Datos

Asegúrese de que su conjunto de prueba cumpla con las siguientes condiciones:

- Que sea lo suficientemente grande como para producir resultados estadísticamente significativos
- Que sea representativo del conjunto de datos. No elija un conjunto que contenga características que sean diferentes del conjunto de datos de entrenamiento.



Training Data

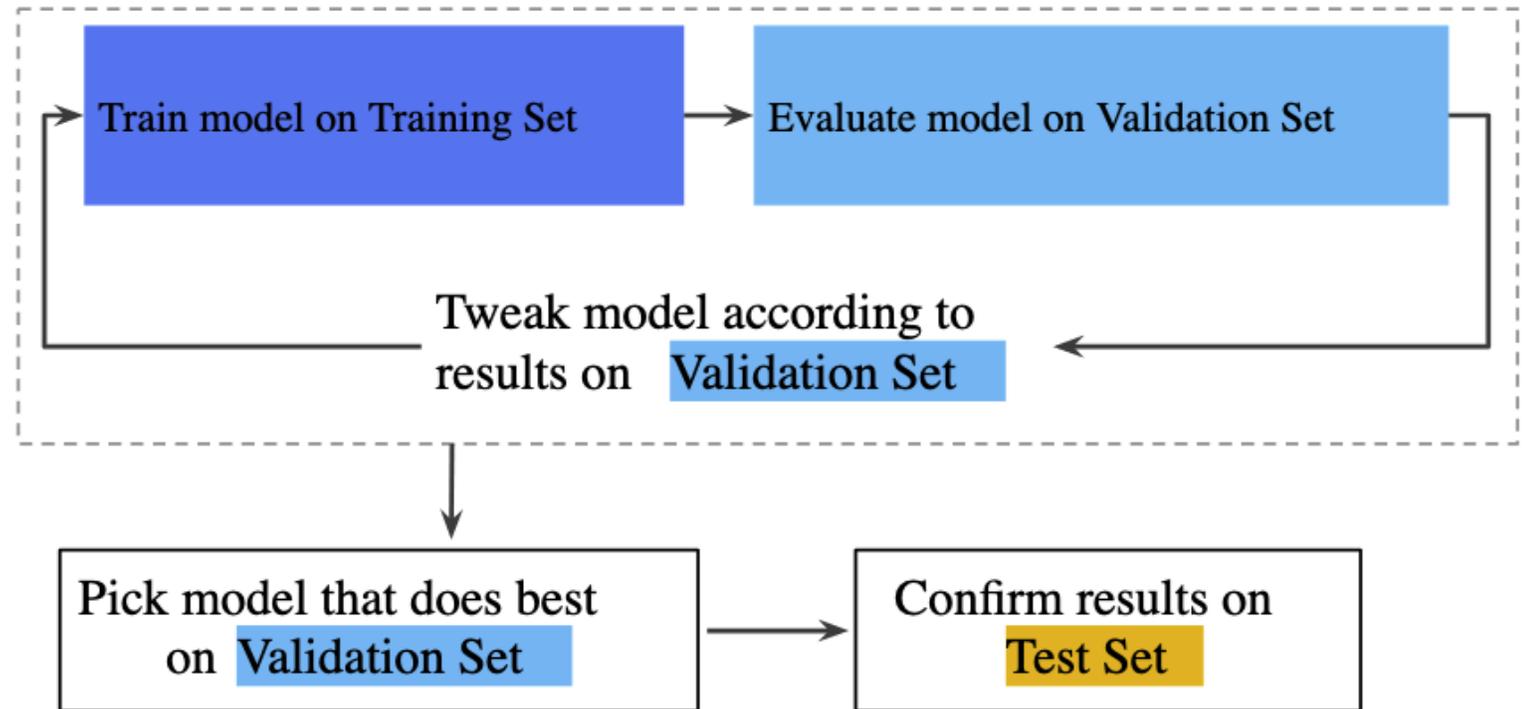


Test Data



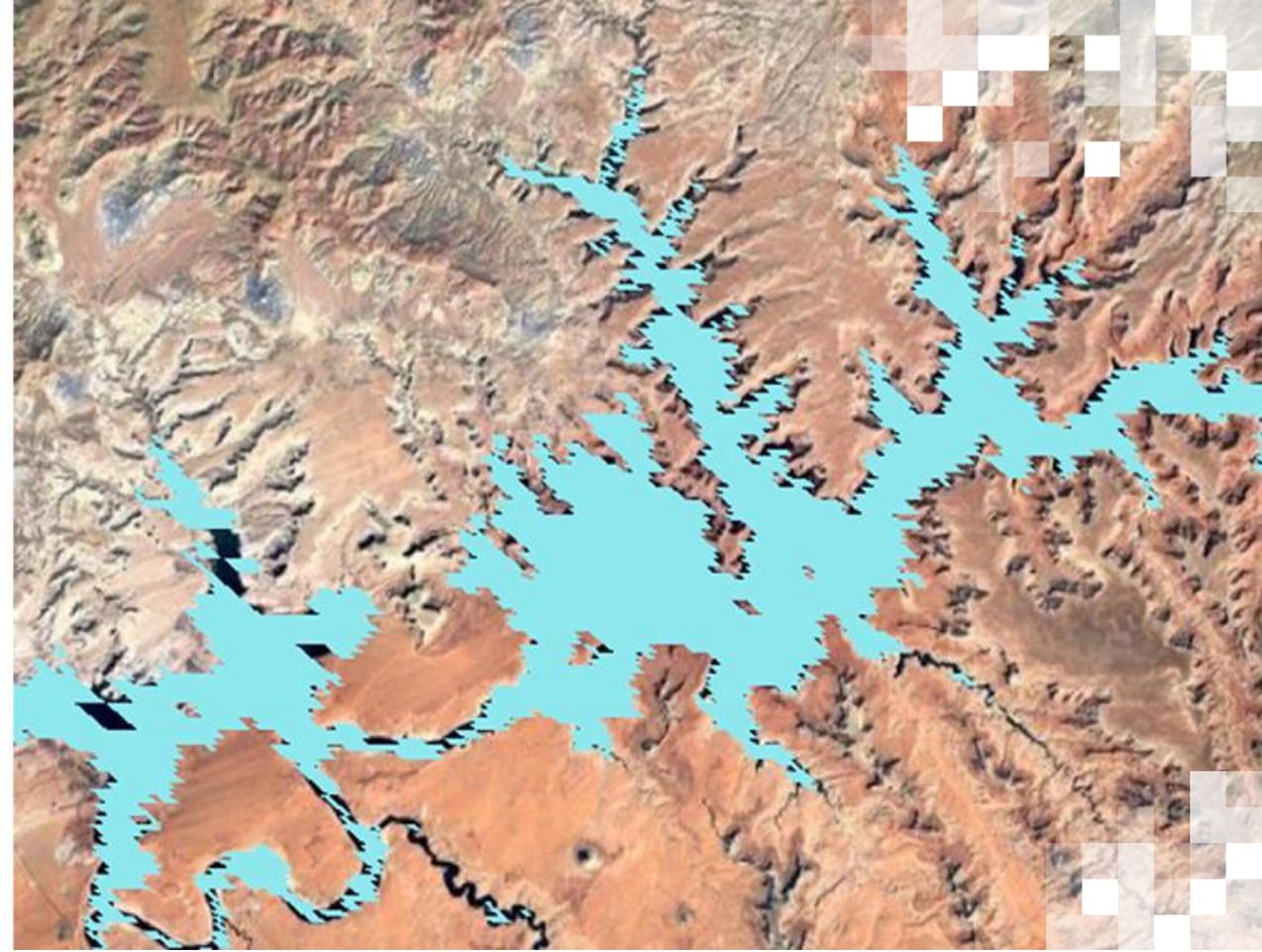
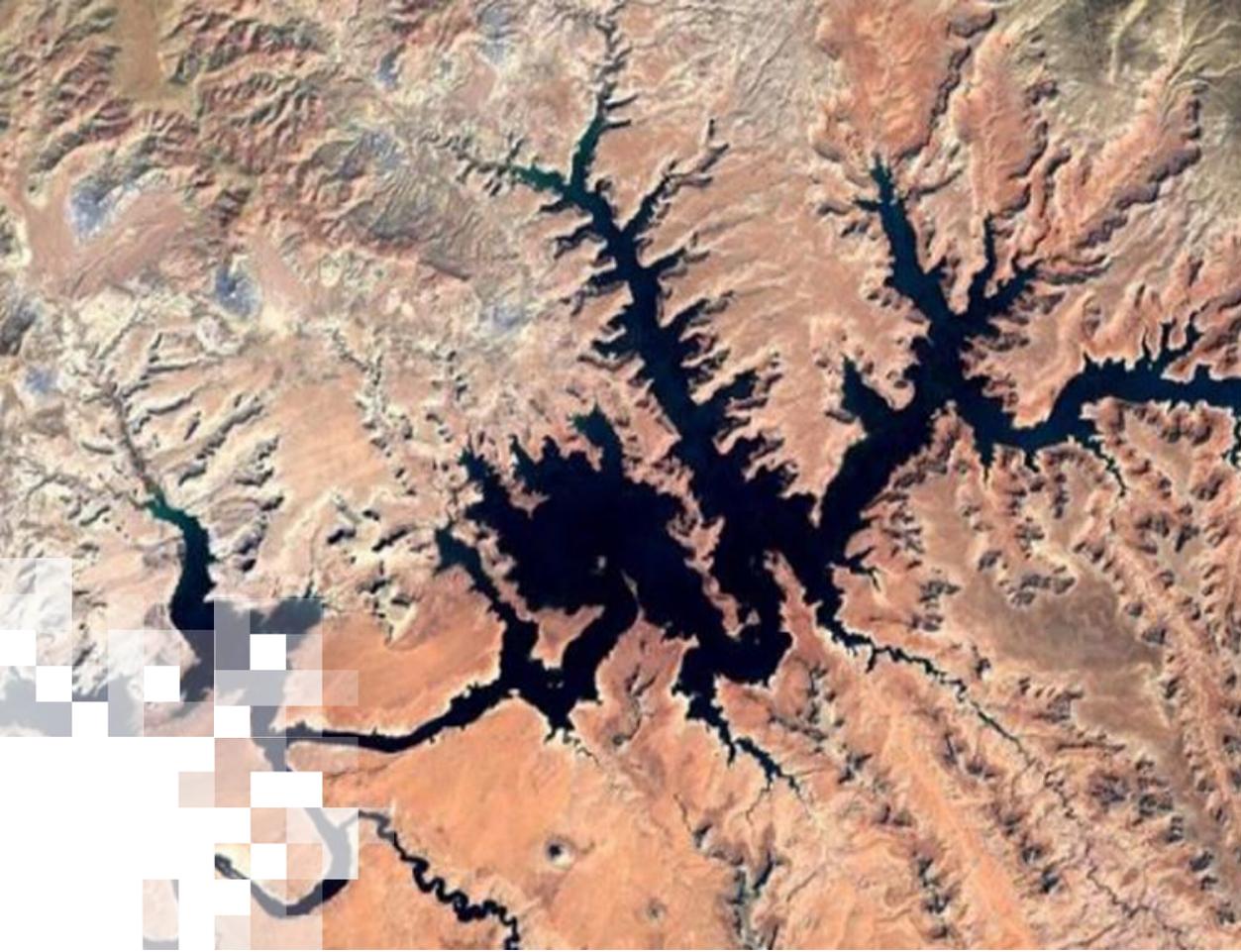
Partición de Datos – ¿Cómo se Alinea en el Flujo de Trabajo de ML?

1. Entrenar el modelo con el conjunto de entrenamiento
2. Usar el conjunto de validación para evaluar los resultados del entrenamiento
3. Usar el conjunto de prueba para confirmar el rendimiento después de que el modelo haya funcionado lo suficientemente bien con el conjunto de validación



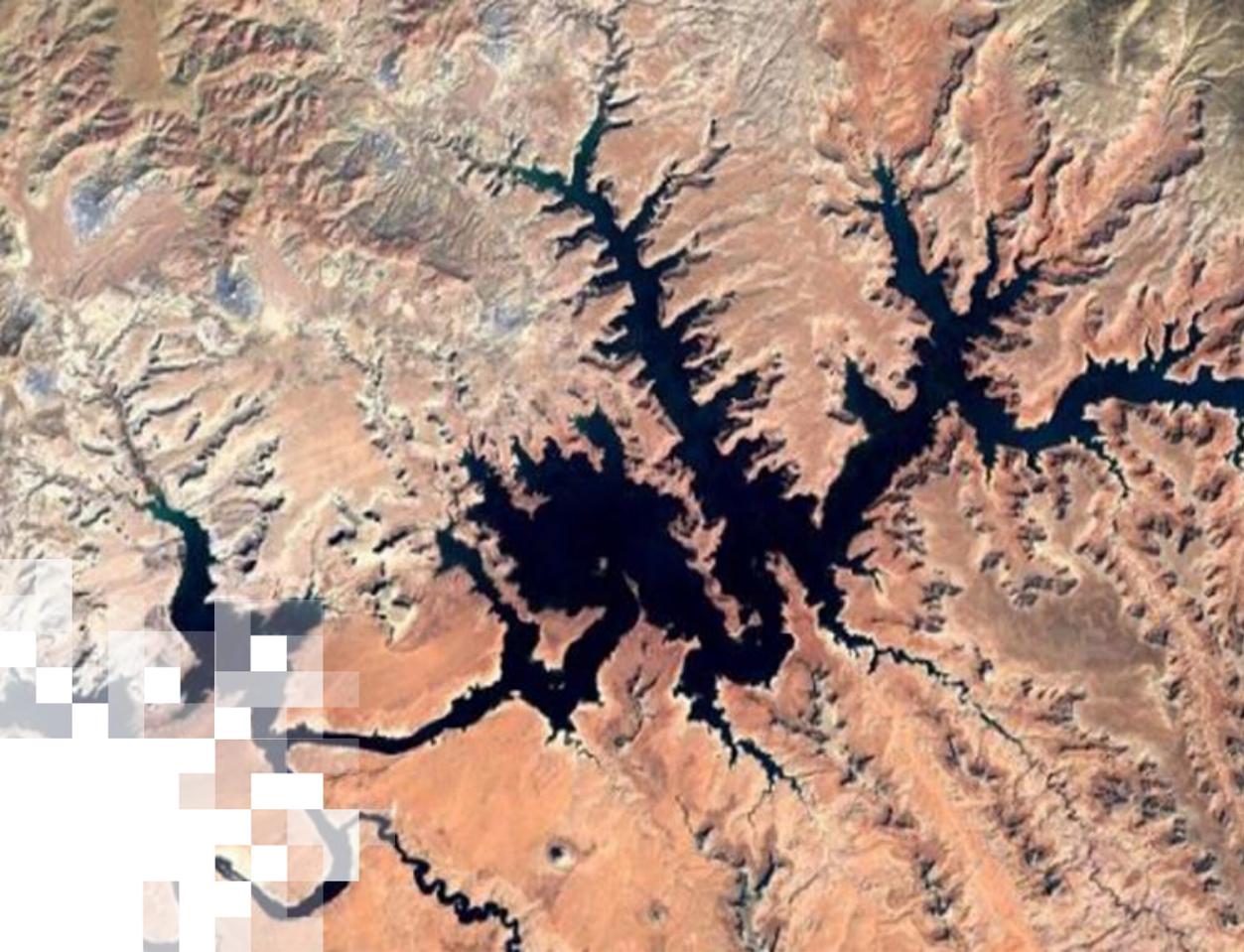
Fuente de la Imagen:
developers.google.com





Ejercicio: Análisis de Datos Exploratorio (Exploratory Data Analysis o EDA) en Google Colab

Formador: Caleb S. Spradlin



Ejercicio: Entrenamiento y Prueba de un Modelo Random Forest en Google Colab

Formador: Jules Kouatchou

Resumen

- Descargar los datos de entrenamiento
- Análisis de datos exploratorio
- Extracción de datos de entrenamiento de un conjunto de datos tabulares
- Extracción de datos de entrenamiento de datos ráster
- Entrenamiento e inferencia del conjunto de datos tabulares y ráster
- Métricas y evaluación de modelos
- Ejercicio Práctico en Jupyter Notebook: Estudio de Caso de la Clasificación del Agua de MODIS



La Próxima Sesión

Parte 3: Ajustes de Modelos, Optimización de Parámetros y Algoritmos de Aprendizaje Automático Adicionales

- Ajustar los modelos
- Optimización de parámetros
- Ejercicio para optimizar un modelo existente
- Resumen general de interpretabilidad y contenido de información de los modelos
- Algoritmos de aprendizaje automático adicionales
- Ejercicio práctico en cuadernos de Jupyter: ***Improvements to MODIS Water Classification Model***



Contactos

- Instructores:
 - Jordan A. Caraballo-Vega: jordan.a.caraballo-vega@nasa.gov
 - Jules Kouatchou: jules.kouatchou-1@nasa.gov
 - Caleb S. Spradlin: caleb.s.spradlin@nasa.gov
 - Jian Li: jian.li@nasa.gov
- Página Web de la Capacitación:
 - <https://appliedsciences.nasa.gov/join-mission/training/spanish/arset-fundamentos-del-aprendizaje-automatico-para-las-ciencias-de-la>
- Página Web de ARSET:
 - <https://appliedsciences.nasa.gov/arset>

Check out our sister programs:



¿Preguntas?

- Por favor escriban sus preguntas en la ventana de “Questions”. Las responderemos en el orden que las recibimos.
- Publicaremos las preguntas y respuestas a la página web de la capacitación después de la conclusión del webinar.



Referencias

- Crankshaw, D. y Gonzalez, J. (2018). Prediction-Serving Systems: What happens when we wish to actually deploy a machine learning model to production?. *Queue*, 16(1), 83-97.
- Elders, A., Carroll, M. L., Neigh, C. S., D'Agostino, A. L., Ksoll, C., Wooten, M. R. y Brown, M. E. (2022). Estimating crop type and yield of small holder fields in Burkina Faso using multi-day Sentinel-2. *Remote Sensing Applications: Society and Environment*, 27, 100820.
- Fleming, S. W., Watson, J. R., Ellenson, A., Cannon, A. J. y Vesselinov, V. C. (2021). Machine learning in Earth and environmental science requires education and research policy reforms. *Nature Geoscience*, 14(12), 878-880.
- Prša, A., Kochoska, A., Conroy, K. E., Eisner, N., Hey, D. R., IJspeert, L., ... y Winn, J. N. (2022). TESS Eclipsing Binary Stars. I. Short-cadence Observations of 4584 Eclipsing Binaries in Sectors 1–26. *The Astrophysical Journal Supplement Series*, 258(1), 16.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., y Carvalhais, N. the National Energy Research Supercomputing Center in Lawrence Berkeley National Laboratory, Berkeley, CA, USA: Deep learning and process understanding for data-driven Earth system science. *Nature*, 566, 195-204.
- Yu, S., y Ma, J. (2021). Deep learning for geophysics: Current and future trends. *Reviews of Geophysics*, 59(3), e2021RG000742.



Contribuidores

- Jordan A. Caraballo-Vega
- Mark L. Carroll
- Jules R. Kouatchou
- Jian Li
- Caleb S. Spradlin
- Brock Blevins
- Melanie Follette-Cook
- Erika Podest
- Brian Powell
- Akiko Elders





¡Gracias!

