

AI.NODO

de colaboración Científico-Industrial
para la investigación y desarrollo
de la Inteligencia Artificial



Centro de
Colaboración
de Alto
Desempeño



UNC

FAMAF

Ministerio de
CIENCIA Y
TECNOLOGÍA



Diplomatura en Ciencia de Datos, Aprendizaje Automático y sus Aplicaciones

Mentorías DiploDatos 2024

Coordinadoras: Carolina Chavero - Yanina Iberra

<https://diplodatos.famaf.unc.edu.ar/>

Pautas Generales

Es responsabilidad del mentor fomentar activamente el trabajo en equipo, incentivando la colaboración entre los estudiantes y promoviendo un ambiente en el que puedan compartir conocimientos, ideas y experiencias para enriquecer el proceso de aprendizaje colectivo

- 3 Trabajos Prácticos (acordes a las [materias obligatorias](#)):
 - TP1- Análisis y Visualización
 - TP2 - Exploración y curación
 - TP3 - Supervisado y/o No supervisado (en caso de otra técnica como PLN, se requerirá una guía más fuerte por parte del mentor.)
- Reuniones periódicas para guiar a los alumnos (tendrán dos grupos de 4 personas cada uno aprox.)
- Respetar las fechas, en especial las de los videos, que no son negociables.
- Disponibilizar TPs y Datasets por medio de un repositorio.
- Hacer devoluciones de los TP (oral, escrito, etc). Las entregas y correcciones de los TPs se realizarán por medio del [aula virtual](#).
- Consultas administrativas: patricia.caceres@unc.edu.ar

Proyectos de Mentorías 2024

Diplomatura en Ciencia de Datos, Aprendizaje Automático y sus Aplicaciones

FAMAF

<https://diplodatos.famaf.unc.edu.ar/>

Mentorías

Las mentorías de la Diplomatura en Ciencia de Datos, Aprendizaje Automático y sus Aplicaciones permiten a los estudiantes desarrollar un proyecto a lo largo de toda la formación, con acompañamiento personalizado por una mentora o mentor y en grupos reducidos.

Coordinadoras:

[Carolina Chavero - carolina.chavero@unc.edu.ar](mailto:carolina.chavero@unc.edu.ar)

[Yanina Iberra - yaninaiberra@gmail.com](mailto:YaninaIberra-yaninaiberra@gmail.com)

Proyectos de mentorías ofrecidos en la Diplodatos 2024

[M01: Análisis de gestión de turnos médicos: las conversaciones son más que palabras](#)

[M02: Predicción de Series Temporales Financieras con Machine Learning](#)

[M03: Descifrando el Futuro Económico: Predicción de la Inflación con Machine Learning](#)

[M04: Análisis y Predicción de Calidad del Aire en una Zona Urbana: Exploración de Datos Multisensor y Modelado Predictivo](#)

[M05: Predicción de Churn en Servicios de Telefonía mediante Modelos de Aprendizaje Supervisado](#)

[M06: Cambio climático y ML: cómo mitigar las emisiones de CO2 mediante la reducción del consumo energético en construcciones edilicias.](#)

[M07: Caracterización y Agrupación de Usuarios en Redes Sociales mediante Análisis de Mensajes](#)

[M08: Conexiones Transparentes: Descubriendo Relaciones Cruciales para la Calidad del Agua en el Río de La Plata](#)

[M09: Desafío EPEC: Optimización del consumo energético, realizando predicciones de demanda.](#)

[M10: Proyecto: Sin bajar la Guardia - Construyendo conocimiento al servicio de la salud.](#)

[M11: Predicción de la actividad eléctrica atmosférica del modelo WRF-ELEC empleando técnicas de Machine Learning](#)

[M12: Predicción del tránsito en la Ciudad de Córdoba, aportes para la construcción de un mapa de ruido](#)

[M13: Minería de texto y modelado de tópicos en foros de ansiedad y depresión de Reddit](#)

[M14: Predicción de Consumo de Electricidad en Industria Metalúrgica de Córdoba](#)

[M15: Exploración y Análisis de Comportamiento de Compras en Plataforma de Cursos Online](#)

[M16: Predicción de Accidentabilidad Laboral para Mejorar la Seguridad en Establecimientos](#)

M08: "Conexiones Transparentes: Descubriendo Relaciones Cruciales para la Calidad del Agua en el Río de La Plata"

Mentora: Maricel Santos - [Linkedin](#)

email: maricel.santos@mi.unc.edu.ar

Descripción y objetivo del proyecto :

En este proyecto se propone estudiar un dataset que contiene información de monitoreo de agua para diferentes periodos y estaciones a lo largo de la costa del Río de La Plata. El objetivo es encontrar las interrelaciones entre variables, así como la sensibilidad de las mismas, además de tratar de encontrar posibles patrones relacionados a la estacionalidad y las características de las poblaciones circundantes a la estación monitoreada. Esto puede ser valioso a la hora de tomar decisiones gubernamentales para evitar y penar la contaminación protegiendo así la calidad de nuestras aguas.

El análisis de agua tiene en cuenta aspectos físicos, químicos, microbiológicos y organolépticos. Es importante reconocer que algunos de estos parámetros de control son más costosos de evaluar, tanto en términos de recursos como de tiempo necesario. En este contexto, resulta particularmente interesante explorar las posibles correlaciones entre estas variables utilizando herramientas del aprendizaje supervisado. Este enfoque permitirá no solo comprender las relaciones intrínsecas entre los diversos aspectos del agua, sino también identificar oportunidades para optimizar los recursos, centrándose en la predicción de variables mediante la observación y monitorización de aquellas menos costosas y que incluso en algunos casos pueden ser medidas en tiempo real.

Con éste proyecto trataremos de responder algunas de las siguientes preguntas:

¿En una estación de monitoreo, todas la variables medidas son importantes para el índice de calidad de agua?

¿Existe relación entre la calidad de agua y el periodo de medición?

¿Existe relación entre la calidad de agua y la población del sitio de monitoreo? y entre las características que tiene esa población?

¿Podemos encontrar relaciones entre las variables medidas en el agua?

¿Es posible utilizar correlaciones de forma tal que algunas variables predigan a otras?

¿Es posible utilizar sólo variables de menor costo de monitoreo para asegurar la calidad de agua?

Datos

El dataset constará de la conjunción de algunos datos públicos del Ministerio de Ambiente y Desarrollo Sostenible (2015-2023). El registro cuenta con información de 42 estaciones, monitoreadas entre una y cuatro veces por año (primavera, verano, otoño. Invierno). Se compilará información del año 2022, año en el cual también fueron registradas más de 200 sustancias químicas medidas en superficie y en sedimento. A su vez, se anexa información sobre la población, características y actividades del municipio al que pertenece el sitio de monitoreo. El dataset resultante se encuentra en el siguiente repositorio <https://github.com/MaricelSantos/Mentoria-Diplodatos-2024>, donde se tiene acceso al dataset y la metadata (o explicación de cada variable).

Fechas importantes:

- **16/05:** Las coordinadoras presentamos las descripciones de las mentorías a los alumnos
- **16-20/05:** abriremos el formulario de inscripción para que los estudiantes elijan sus mentorías
- **23-24/05:** notificación a mentores y alumnos de cómo quedaron conformados los grupos de las mentorías
- **Del 27/05 al 31/05:** primera reunión de cada mentor con sus grupos para presentar el dataset y el TP1
- **01/07 Mentorías** - Los alumnos entregan vía aula virtual el **práctico 1**. Inmediatamente los mentores le presentan el **Práctico 2**. Una vez evaluado el práctico, hacen la devolución y cargan en el aula virtual “Aprobado/No Aprobado/Revisar”
- **29/07** - Los alumnos entregan el **práctico 2**: se hace la devolución pertinente y se les da las pautas para la presentación del video intermedio que resume el P1 y el P2, duración <10', participen todos en la medida de lo posible, destaquen el “Storytelling”. Sin bloques de código!
- **12/08** - Los alumnos entregan del video de presentación intermedia del proyecto (**video - presentación 1**) . Los mentores les presentan el Práctico 3.
- **16/09** - Los alumnos entregan el **práctico 3**, se da la devolución y dan el visto bueno para presentar el video final.
- **30/09** - Entrega del video de presentación final de la mentoría (**video - presentación 2**)
- **15/11 y 16/11** - Jornadas de presentación de videos Finales.

(todas estas instancias son de modalidad virtual)