

Maestría en Ciencia de Datos y Aprendizaje Automático

Propuesta Tesis Maestría

“Gestión de datos faltantes en monitoreo ambiental casi en tiempo real”

El problema de los datos faltantes ocurre frecuentemente en el sector hidráulico-ambiental. En la última década, se han adoptado técnicas de aprendizaje automático para reconstruir series temporales ambientales [1].

En el marco del proyecto FSDA_1_2018_1_153967, financiado por ANII, con título “Evaluación temporal y espacial del impacto del cambio de cobertura del suelo sobre la calidad del agua: Cuenca del río Santa Lucía como cuenca piloto” [2,3], se desarrolló una metodología híbrida (con técnicas de aprendizaje automático alimentadas de información físicamente basada) para la imputación de diferentes variables ambientales.

Basándonos en dicha metodología desarrollada en la cuenca piloto del río Santa Lucía, esta propuesta de tesis tiene los siguientes objetivos:

- Extensión y validación de la metodología híbrida de imputación de datos meteorológicos, hidrológicos y de calidad de aguas a nivel nacional.
- Creación de una base de datos multivariados que se auto-actualice a partir de las fuentes de datos nacionales disponibles.
- Diseño y desarrollo de una aplicación informática que ejecute la metodología de imputación y permita a los usuarios visualizar la base de datos multivariados imputada y actualizada.

El estudiante trabajará en un grupo mixto multidisciplinario que integra investigadores del Instituto de Mecánica de los Fluidos e Ingeniería Ambiental (IMFIA) y del Instituto de Computación (INCO), de la Facultad de Ingeniería (FIng), Universidad de la República (UdelaR).

Se busca un estudiante con un interés y motivación en datos ambientales y en el desarrollo de aplicaciones donde el cuidado de los aspectos gráficos, estéticos e interactivos esté presente.

Referencias:

[1] Zhang, Y.; Thorburn, P.J. Handling missing data in near real-time environmental monitoring: A system and a review of selected methods. *Future Generation Computer Systems* 2022, 128, 63–72.

[2] Gorgoglione, A.; Castro, A.; Rodríguez Núñez, R.; Chreties, C.; Fossati, M.; Pastorini, M.; Etcheverry, L. Evaluación temporal y espacial del impacto del cambio de cobertura del suelo sobre la calidad del agua: Cuenca del río Santa Lucía como cuenca piloto. *Udelar.FI.*, 2021.

[3] Rodríguez, R.; Pastorini, M.; Etcheverry, L.; Chreties, C.; Fossati, M.; Castro, A.; Gorgoglione, A. Water-Quality Data Imputation with a High Percentage of Missing Values: A Machine Learning Approach. *Sustainability* 2021, 13, 6318.

Contactos:

Angela Gorgoglione – agorgoglione@fing.edu.uy

Alberto Castro - acastro@fing.edu.uy