

MODELOS GENERATIVOS INTERPRETABLES PARA DISEÑO *DE NOVO* DE MOLÉCULAS

Trinidad Crozes^{1,2}, Vir Sabando¹, Axel J. Soto^{1,2}, Ignacio Ponzoni^{1,2}

¹ Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur, Argentina

² Instituto de Ciencias e Ingeniería de la Computación, CONICET-UNS, Argentina

ip@cs.uns.edu.ar



CONTEXTO

Proyecto de investigación:

“Aprendizaje automático y analítica visual para el uso de modelos generativos y predictivos aplicados a quimioinformática y el procesamiento de lenguaje natural”.

Dirigido por los Dres. Axel J. Soto e Ignacio Ponzoni.

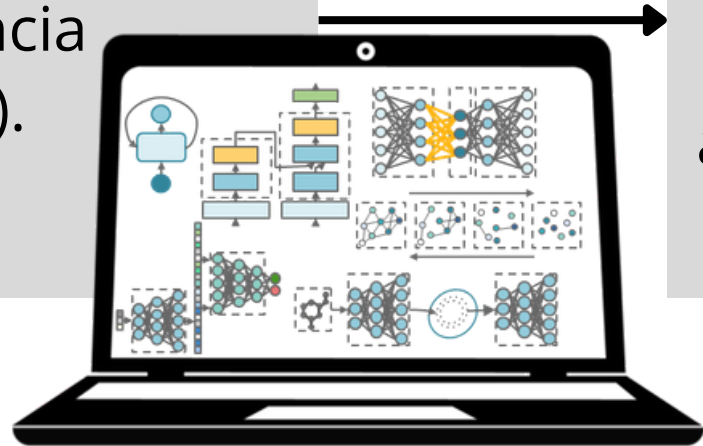
El proyecto es coordinado y financiado por la Universidad Nacional del Sur, con la contribución del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) quien financia la beca doctoral de la Lic. Crozes.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

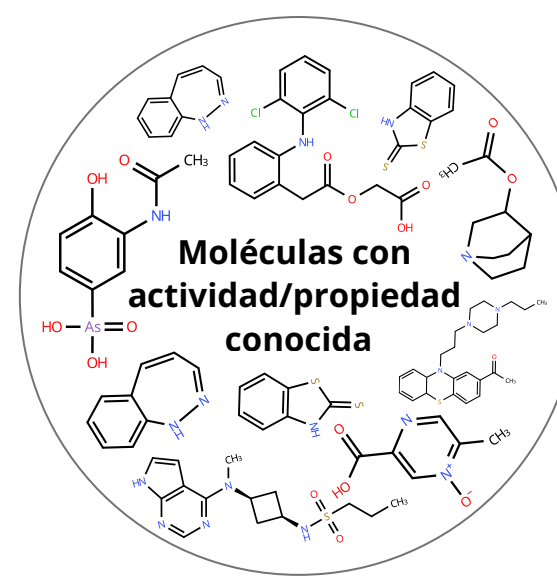
- 1 ¿Qué tipo de algoritmo generativo es más adecuado para el modelado molecular y cómo impactan las distintas representaciones moleculares?
- 2 ¿Cuál estrategia de modelado QSAR (*Quantitative Structure-Activity Relationship*) logra una mayor efectividad al integrar la evaluación de varias propiedades fisicoquímicas y de actividad biológica para validar el modelo generativo?
- 3 ¿Cómo una estrategia de analítica visual puede mejorar la incorporación del conocimiento experto y brindar apoyo a la interpretación de los resultados?
- 4 ¿Cómo se puede validar el desempeño de la propuesta en el diseño de quimiotecas virtuales para proyectos de descubrimiento de fármacos?

RESULTADOS OBTENIDOS/ESPERADOS

Revisión de técnicas para **generación de moléculas** basadas en inteligencia artificial (IA).



Diseñar **modelos QSAR** para evaluar propiedades fisicoquímicas y de actividad biológica de interés.



MODELO QSAR basado en técnicas de aprendizaje automático

Molécula con actividad/propiedad desconocida → **Predicción de la actividad/propiedad**

Integrar en una metodología basada en **aprendizaje por refuerzo** a los modelos generativos y QSAR más prometedores.

Diseñar una **estrategia de IA explicable** para evaluar e interpretar los resultados obtenidos.

Diseñar una **estrategia de analítica visual interactiva** que brinde soporte a las estrategias de IA interpretable propuestas.

Integrar estas técnicas en una **herramienta computacional** que facilite a expertos en química medicinal hacer **diseño de novo de compuestos**.

FORMACIÓN DE RECURSOS HUMANOS

El doctorado de la Lic. Trinidad Crozes se desarrolla en el marco de la presente línea de investigación, radicada en el grupo BioChemTICs del Instituto de Ciencias e Ingeniería de la Computación (CONICET-UNS).

Temática central: Diseño de metodologías de aprendizaje automático orientadas a la inferencia y análisis de redes jerárquicas a partir de datos de expresión génica, en el ámbito de la bioinformática. Disciplina: Informática y Comunicaciones, con aplicación en ciencias exactas y naturales.