**Información Relevante Proyecto de Grado**

**Información del Estudiante**

Germán de Jesús Torres Arroyave

[german.torres@udea.edu.co](mailto:german.torres@udea.edu.co)

1001812815

**Información del Profesor Director**

Hernan David Salinas Jimenez

[hernan.salinas@udea.edu.co](mailto:hernan.salinas@udea.edu.co)

Grupo de Magnetismo y Simulación G+

**Información del Profesor co-Director**

Cesar Augusto Barrero Meneses

[cesar.barrero@udea.edu.co](mailto:cesar.barrero@udea.edu.co)

Grupo de Estado Sólido

**Título del Proyecto**

**Physics Informed Machine Learning: Una Perspectiva en el Análisis Cinético de Adsorción de Contaminantes en el Agua.**

**Justificación del Proyecto**

El tratamiento de aguas contaminadas por metales pesados es una de las problemáticas más críticas a nivel global, debido a los graves efectos que estos contaminantes tienen sobre la salud humana y el medio ambiente [1]. La adsorción ha demostrado ser uno de los métodos más eficaces y costo-efectivos para la remoción de estos contaminantes. Sin embargo, la optimización de este proceso a escala industrial enfrenta desafíos significativos, como la variabilidad en las condiciones del agua y la necesidad de adsorbentes eficientes y sostenibles [2-5].

En respuesta a estos desafíos, los recientes avances en inteligencia artificial, particularmente en redes neuronales artificiales (ANN), han permitido modelar y predecir con alta precisión la capacidad de adsorción de diferentes adsorbentes bajo diversas condiciones operativas. Estos modelos no solo mejoran la eficiencia del proceso, sino que también reducen los costos operativos y minimizan el uso de químicos, lo que es crucial para el desarrollo de tecnologías de tratamiento de agua más sostenibles [3-5].

Sin embargo, los modelos puramente basados en datos, aunque poderosos, presentan limitaciones en situaciones donde los datos son escasos, ruidosos o incompletos. En muchos casos, estos modelos pueden generar predicciones que no respetan las leyes físicas subyacentes, lo que compromete su fiabilidad​ [6]. En este contexto, las técnicas de Physics Informed Machine Learning (PIML) emergen como una solución prometedora.

Las técnicas de PIML integran las leyes físicas fundamentales, como las ecuaciones diferenciales que gobiernan los procesos de adsorción, dentro de los modelos de machine learning. Esta integración no solo mejora la precisión de las predicciones al aprovechar tanto los datos como los principios físicos, sino que también proporciona mayor robustez frente a datos imperfectos o incompletos [6]. Esta capacidad es particularmente ventajosa en el estudio de la cinética de adsorción de metales pesados, donde los modelos basados en PIML pueden generar predicciones físicamente consistentes incluso en condiciones de incertidumbre experimental.

**Objetivo General**

Desarrollar e implementar técnicas de Physics Informed Machine Learning (PIML) para modelar, analizar y predecir la cinética de adsorción de contaminantes metálicos en el agua, integrando datos experimentales con modelos basados en ecuaciones cinéticas que describen el proceso.

**Objetivos Específicos**

1. Estudiar e implementar modelos de PIML aplicados a la solución de ecuaciones diferenciales e interpretación de datos experimentales.
2. Ordenar bases de datos e implementar algoritmos modulares y flexibles de PIML para entender la cinética de adsorción de contaminantes metálicos en el agua
3. Entrenar, validar y analizar los modelos construidos para la cinética asociada al proceso de adsorsión a partir de los datos experimentales realizando comparación con algunos modelos existentes, como: el de primer orden, segundo orden, Avrami, logística, entre otros.

**Metodología**

Para cumplir todos los objetivos será necesario realizar una revisión constante de literatura y fundamentos teóricos, además de reportar y documentar todo el proceso de exploración, desarrollo de modelos, resultados y conclusiones. En particular para cumplir cada objetivo seguiremos lo siguiente:

**Objetivo específico 1**: Para desarrollar modelos robustos en la cinética de adsorción de contaminantes metálicos en el agua, es esencial revisar la literatura y construir modelos PIML validados. Este enfoque integra principios físicos con machine learning, mejorando la precisión, reduciendo la necesidad de datos y permitiendo una mejor generalización. Además, facilita la interpretación y es eficaz en la resolución de ecuaciones diferenciales complejas, combinando datos experimentales con teorías físicas.

Los detalles de la metodología para este objetivo se presentan a continuación:

1.1. Revisión de literatura y fundamentos teóricos: Investigar artículos científicos, libros y publicaciones recientes sobre PIML, identificando métodos, algoritmos y arquitecturas de redes neuronales, así como aplicaciones exitosas en la solución de ecuaciones diferenciales y análisis de datos experimentales. Repasar conceptos de machine learning, teoría de ecuaciones diferenciales y métodos numéricos, comprendiendo cómo integrar conocimiento físico en modelos de machine learning. En particular, se estudiará el artículo de Karniadakis et al. en Nature, entre otros autores, a lo largo del proyecto.

1.2. Construcción y validación de modelos: Además de la revisión constante de bibliografía, será necesario construir y validar modelos simples que puedan ser implementados en el estudio de la cinética asociado a la formación de cristales. Para ello, será necesario estudiar situaciones simples que involucren soluciones experimentales y teóricas en las que podemos analizar el comportamiento de modelos de PIML aplicados a distintos tipos de datos, para profundizar en el funcionamiento y el entendimiento de los procesos que se llevan a cabo dentro de las PIML, a saber: péndulo simple, oscilador armónico etc.

**Objetivo específico 2.** Con los códigos construidos en el apartado anterior, y las investigaciones de estudios recientes sobre la cinética de adsorción de contaminantes metálicos en el agua, realizaremos la implementación de nuevos códigos. Detalles de los métodos para cumplir este objetivo son dados a continuación:

**2.1. Revisión de literatura y fundamentos teóricos.** Investigaremos estudios recientes sobre la cinética de adsorción de contaminantes metálicos en el agua y los factores que influyen en este proceso, tales como el uso de nanoadsorbentes magnéticos y técnicas de caracterización experimental. Además, revisaremos los conceptos fundamentales relacionados con los procesos de adsorción y las ecuaciones cinéticas que modelan dichos fenómenos.

**2.2. Obtención, exploración, limpieza y análisis de los datos.** Con los datos experimentales proporcionados por el grupo de investigación encargado del proyecto *"FABRICACIÓN, CARACTERIZACIÓN Y USO DE NANOADSORBENTES MAGNÉTICOS BASADOS EN AKAGANEITAS Y FERRITAS FUNCIONALIZADAS CON EDTA PARA ADSORBER METALES PESADOS EN AGUA"*, construiremos una base de datos que incluirá curvas de concentración de metales pesados adsorbidos en función del tiempo bajo distintas condiciones experimentales. Además, evaluaremos la posibilidad de integrar otras fuentes de datos externas para enriquecer nuestro análisis. Para asegurar la calidad y consistencia de los datos, se llevará a cabo una exhaustiva etapa de exploración y limpieza de los mismos.

**2.3.** **Desarrollo de código fuente.** Con la base de datos definida y los algoritmos desarrollados en el Objetivo 1, implementaremos códigos en Python utilizando PyTorch para modelos PIML. Estos serán modulares y flexibles en: i) la variación de arquitecturas de redes neuronales, ii) la lectura de datos experimentales, y iii) la implementación de ecuaciones diferenciales cinéticas. Incorporaremos restricciones físicas específicas del proceso de adsorción, utilizando tanto los datos experimentales como las ecuaciones diferenciales derivadas de los modelos cinéticos del proceso.

**Objetivo específico 3.** Para cumplir este objetivo, será necesaria una colaboración constante con el grupo de investigación a cargo del proyecto "FABRICACIÓN, CARACTERIZACIÓN Y USO DE NANOADSORBENTES MAGNÉTICOS BASADOS EN AKAGANEITAS Y FERRITAS FUNCIONALIZADAS CON EDTA PARA ADSORBER METALES PESADOS EN AGUA" para validar y entender la naturaleza de los datos experimentales. Con los códigos construidos en el apartado anterior y el análisis de estudios recientes sobre la cinética de adsorción, enfocados en modelos cinéticos existentes, seguiremos la siguiente estrategia:

**3.1 Entrenamiento.** Realizaremos el entrenamiento de los modelos PIML; para ello, será fundamental estudiar las funciones de pérdida que mejor integren los datos experimentales y las ecuaciones cinéticas. La optimización de la función de pérdida será clave para asegurar la coherencia física del modelo, ajustando hiperparámetros como el número de capas en la red neuronal, los pesos de las ecuaciones diferenciales y los pesos asignados a los datos experimentales, entre otros.

**3.2 Validación, comparación de técnicas e interpretación de resultados.** Esta etapa será crucial en el desarrollo del proyecto. Aquí, seleccionaremos modelos cinéticos existentes para compararlos con los resultados obtenidos mediante la implementación de técnicas de PIML. Utilizaremos métodos estadísticos para evaluar la precisión y robustez de cada modelo. Además, desarrollaremos visualizaciones que faciliten la comparación de los resultados, mostrando gráficamente la correspondencia entre los modelos teóricos y los datos experimentales.

**3.3 Interpretación de Resultados.** Analizaremos los resultados obtenidos, evaluando la precisión y robustez de los modelos PIML en la predicción de la cinética de adsorción de contaminantes metálicos. Desarrollaremos visualizaciones que faciliten la interpretación y comparación de los resultados entre diferentes condiciones experimentales, proporcionando una mejor comprensión de los procesos de adsorción.

**Antecedentes**

Al momento de presentar la propuesta de este proyecto, he completado con éxito un curso electivo del pregrado de Física titulado 'Aprendizaje Estadístico.' Este curso se enfoca en el análisis de datos desde una perspectiva estadística y abarca la aplicación de modelos de regresión, así como el desarrollo y evaluación de modelos de aprendizaje automático supervisados y no supervisados, con énfasis en la capacidad de estos modelos para identificar patrones, hacer predicciones y optimizar resultados en diversos conjuntos de datos.

Además, como parte de mi experiencia en este curso, desarrollé un proyecto final titulado *'Optimización de un modelo de red neuronal convolucional 1D para identificar el sistema cristalino de compuestos inorgánicos binarios.'* Este proyecto consistió en diseñar y optimizar un modelo de clasificación capaz de identificar con alta precisión el sistema cristalino de compuestos binarios inorgánicos a partir de patrones de difracción de rayos X, demostrando la aplicación efectiva de técnicas avanzadas de aprendizaje automático en el campo de la Física.

**Cronograma**

**Semanas 1-4**

* Revisión de literatura y fundamentos teóricos
* Construcción y validación de modelos de exploración

**Semanas 5-9**

* Obtención, exploración, limpieza y análisis de los datos.
* Desarrollo de modelos y algoritmos: modulares y escalables

**Semanas 10-16**

* Entrenamiento de los modelos y cálculos de las funciones de coste
* Validación y retroalimentación de los modelos
* Correcciones y mejoras de los modelos
* Comparación de técnicas e interpretación de resultados

**Requerimientos específicos y presupuesto**

**Computador Personal:** Se cuenta con un computador personal adecuado para el desarrollo del proyecto. El equipo cumple con los requisitos necesarios para ejecutar el software y realizar las tareas de análisis y procesamiento de datos.

No es necesario incluir este ítem en el presupuesto, ya que se dispone de un equipo propio con las especificaciones requeridas.

**Referencias Bibliográficas**

[1] Sankhla, M. S., Kumari, M., Nandan, M., Kumar, R., & Agrawal, P. (2016). Heavy metals contamination in water and their hazardous effect on human health: A review. International Journal of Current Microbiology and Applied Sciences, 5(10), 759-766. http://dx.doi.org/10.20546/ijcmas.2016.510.082

[2] Rashid, R., Shafiq, I., Akhter, P., Iqbal, M. J., & Hussain, M. (2021). A state-of-the-art review on wastewater treatment techniques: The effectiveness of adsorption method. Environmental Science and Pollution Research, 28(9), 9050-9066. https://doi.org/10.1007/s11356-021-12395-x

[3] Karniadakis, G.E., et al. (2021). Physics-Informed Machine Learning: Theory and Applications. Nature Reviews Physics, 3(6), 422-440. https://doi.org/10.1038/s42254-021-00314-5&#8203;:contentReference[oaicite:1]{index=1}

[4] Khan, A.A., et al. (2020). Adsorption of Heavy Metals Using Nanomaterials: Recent Advances and Future Prospects. Molecules, 25(15), 3263. https://doi.org/10.3390/molecules25153263

[5] Kumari, S., Chowdhry, J., Garg, M.C. (2024). Applications of Artificial Intelligence in Water Treatment and Adsorption Processes. Journal of Environmental Management, 351, 119968. https://doi.org/10.1016/j.jenvman.2023.119968&#8203;:contentReference[oaicite:2]{index=2}

[6] Bardiji, N., Ziat, K., Naji, A., & Saidi, M. (2020). Machine Learning Models for Adsorption Kinetics and Equilibrium: A Case Study with Heavy Metals. ACS Omega, 5, 5105-5115. https://doi.org/10.1021/acsomega.9b04088&#8203;:contentReference[oaicite:0]{index=0}