

Tipo de artículo: Artículo original

# Diseño de un sistema para evaluar la eficacia del índice en la predicción de propiedades fisicoquímicas y biológicas de compuestos químicos

## *Design of a system to evaluate the effectiveness of the index in predicting physicochemical and biological properties of chemical compounds*

Claudia M. Cardoso Gallardo<sup>1\*</sup> , <https://orcid.org/0009-0009-9974-4715>

María Teresa de la Cruz Padilla<sup>2</sup> , <https://orcid.org/0009-0004-5931-8504>

<sup>1</sup> Universidad de las Ciencias Informáticas. Cuba. Correo electrónico: [claudiamcg@estudiantes.uci.cu](mailto:claudiamcg@estudiantes.uci.cu)

<sup>2</sup> Universidad de las Ciencias Informáticas. Cuba. Correo electrónico: [mteresadlcp@uci.cu](mailto:mteresadlcp@uci.cu)

\* Autor para correspondencia: [mteresadlcp@uci.cu](mailto:mteresadlcp@uci.cu)

### Resumen

El presente estudio se centra en el diseño de un sistema destinado a evaluar la eficacia del Índice de Suma Inversa (ISI) ponderado en la predicción de propiedades fisicoquímicas y biológicas de compuestos químicos. Para ello, se comparará la capacidad predictiva del ISI ponderado con la de descriptores tradicionales en una muestra representativa de moléculas. Se emplearán tanto métodos teóricos como empíricos para desarrollar y evaluar el desempeño del ISI ponderado en relación con las propiedades fisicoquímicas. Además, se aplicarán técnicas estadísticas avanzadas para analizar la correlación entre los valores obtenidos a partir de los descriptores tradicionales y los valores predichos por el ISI ponderado. Se espera que los resultados obtenidos no solo contribuyan al avance de la quimioinformática y la teoría de grafos, sino que también proporcionen una comprensión más profunda de la relación entre la estructura molecular y las propiedades observadas en los compuestos químicos. Este conocimiento podría tener importantes implicaciones en diversos campos, desde el diseño de fármacos hasta la ingeniería de materiales, ampliando así las aplicaciones prácticas en la ciencia de materiales y la biomedicina.

**Palabras clave:** Diseño de un sistema; evaluar la eficacia; índice en la predicción; propiedades fisicoquímicas y biológicas

### Abstract

*The present study focuses on the design of a system aimed at evaluating the effectiveness of the weighted Inverse Sum Index (ISI) in predicting physicochemical and biological properties of chemical compounds. To do this, the predictive capacity of the weighted ISI will be compared with that of traditional descriptors in a representative sample of molecules. Both theoretical and empirical methods will be employed to develop and evaluate the performance of the weighted ISI in relation to physicochemical properties. In addition, advanced statistical techniques will be applied to analyze the correlation between the values obtained from traditional descriptors and the values predicted by the weighted ISI. The results obtained are expected to not only contribute to the advancement of cheminformatics and graph theory, but also provide a deeper understanding of the relationship between molecular structure and the properties observed in chemical compounds. This knowledge could have important implications in diverse fields, from drug design to materials engineering, thus expanding practical applications in materials science and biomedicine.*

**Keywords:** System design; evaluate effectiveness; index in prediction; physicochemical and biological properties



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional**  
(CC BY 4.0)

**Recibido: 08/02/2024**  
**Aceptado: 24/04/2024**  
**En línea: 01/05/2024**

## Introducción

La quimioinformática emerge como una poderosa herramienta derivada de la fusión entre recursos informáticos y datos químicos, desempeñando un papel crucial en el manejo, visualización y análisis sistemático de la información molecular (Alves et al., 2018). Dentro de este vasto campo, el estudio del Índice de Suma Inversa (ISI) ponderado junto con propiedades fisicoquímicas de los índices topológicos ha ganado prominencia como un descriptor molecular fundamental en la química, particularmente en el desarrollo de relaciones cuantitativas estructura-actividad (QSAR) (Trujillo, 2011).

El presente estudio se centra en el diseño de un sistema destinado a evaluar la eficacia del ISI ponderado en la predicción de propiedades fisicoquímicas y biológicas de compuestos químicos. Una interrogante crucial que guía esta investigación es si el ISI ponderado supera en precisión y efectividad a los descriptores tradicionales en la predicción de propiedades de compuestos químicos, así como si está intrínsecamente correlacionado con las propiedades reales de dichos compuestos (Borroto et al., 2012).

Para abordar esta cuestión, se propone la implementación y comparación del ISI ponderado con descriptores tradicionales en una muestra representativa de moléculas. Esta investigación empleará tanto métodos teóricos como empíricos para desarrollar y evaluar la eficacia del ISI ponderado en relación con las propiedades fisicoquímicas (Torres et al., 2005).

Además, se utilizarán técnicas estadísticas avanzadas para analizar la correlación entre los valores obtenidos a partir de los descriptores tradicionales y los valores predichos por el ISI ponderado. Este enfoque metodológico integral permitirá una evaluación robusta y comparativa de la capacidad predictiva de ambos métodos (Armijos-Jaramillo et al., 2020).

Este estudio no solo contribuirá al avance de la quimioinformática y la teoría de grafos, sino que también arrojará luz sobre la relación entre la estructura molecular y las propiedades observadas en los compuestos químicos (Medina-Franco et al., 2015). Esta comprensión más profunda podría tener importantes implicaciones en diversos campos, desde el diseño de fármacos hasta la ingeniería de materiales, ampliando así las fronteras del conocimiento y la aplicación práctica en la ciencia de materiales y la biomedicina.



## Materiales y métodos

La metodología XP (Programación Extrema) surge en 1996, cuando Kent Beck publicó su libro "Extreme Programming Explained". Desde entonces, ha experimentado varias versiones y mejoras a lo largo de los años. Las versiones más destacadas son XP 1.0, XP 2.0 y XP 3.0, cada una con ajustes y refinamientos en sus prácticas y enfoques (Campos & Martínez, 2015).

A partir de los criterios de selección del proyecto se presenta como resultado la representación de los criterios para la selección del enfoque ver figura 1.

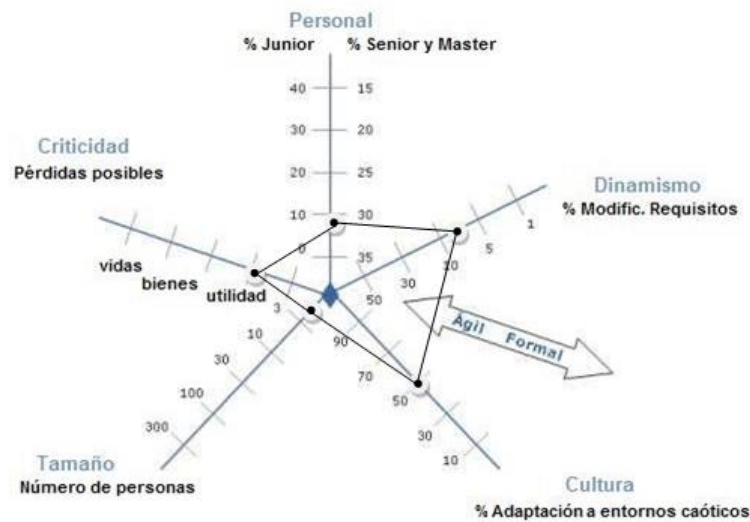


Figura 1. Representación de los criterios para la selección del enfoque.

Algunas de las prácticas más comunes en XP incluyen la programación en parejas, pruebas unitarias continuas, integración continua, desarrollo orientado a pruebas (TDD), diseño simple, refactorización, reuniones rápidas diarias (stand-up meetings) y entregas frecuentes.

**Roles:** Los roles principales en XP son el Cliente (representante del cliente o usuario final), Programador (responsable de escribir el código), Tester (encargado de realizar pruebas) y Entrenador (facilitador que ayuda al equipo a aplicar las prácticas de XP de manera efectiva).

**Actividades definidas (Restricciones vs flexibilidad):** XP define un conjunto de prácticas y valores fundamentales que deben ser seguidos de manera estricta, pero también promueve la flexibilidad para adaptarse a los cambios y necesidades específicas de cada proyecto.



Tamaño del equipo de trabajo: Idealmente, un equipo de trabajo en XP consta de 10 personas o menos. Esto permite una comunicación más efectiva y una toma de decisiones más ágil.

Personal: Se espera que los miembros del equipo sean altamente colaborativos, autoorganizados y estén comprometidos con la entrega de valor al cliente.

Experiencia del equipo: Aunque no es un requisito absoluto, se valora la experiencia técnica y la familiaridad con las prácticas ágiles por parte de los miembros del equipo.

Críticidad: La metodología XP es crítica para proyectos en los que se requiere adaptación rápida a cambios en los requisitos del cliente y en los que la retroalimentación temprana es esencial para el éxito del proyecto.

Planificación: XP promueve la planificación incremental y flexible, con un enfoque en la entrega continua de funcionalidades que agreguen valor al cliente.

Cantidad de iteraciones recomendadas: No hay un número fijo de iteraciones recomendadas en XP, ya que esto puede variar según el proyecto. Sin embargo, se enfatiza la entrega temprana y frecuente de incrementos de software.

Duración recomendada de iteraciones: Las iteraciones en XP suelen tener una duración corta, generalmente entre 1 y 3 semanas. Esto permite una rápida retroalimentación y adaptación a los cambios.

## Resultados y discusión

Para el diseño del sistema propuesto se definieron los siguientes requisitos funcionales

RF1: Procesar la solicitud

RF2: Analizar la solicitud de la molécula

RF3: Calcular los descriptores moleculares

RF4: Devolver los resultados

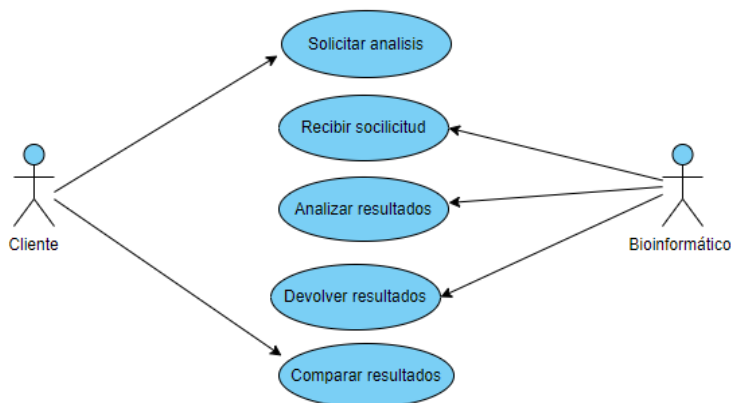
Requisitos No Funcionales:

RNF1: Soportar varias solicitudes al mismo tiempo

RNF2: Tiempo de respuesta máximo de un minuto

Diagrama de casos de uso del negocio





**Figura 2.** Diagrama de Caso de Uso del Negocio

**Tabla 1.** Descripción textual de cada caso de uso del negocio.

<b>Caso de Uso del Negocio</b>	Análisis de Moléculas
<b>Actores</b>	Cliente Bioinformático
<b>Resumen</b>	El cliente realiza una solicitud para el análisis de una molécula; el bioinformático recibe dicha solicitud y procede a realizar el análisis, entrega los resultados y el cliente los compara con otros resultados obtenidos de estudios previos.
<b>Casos de Uso relacionados</b>	-
<b>Acción del actor Cliente</b>	<b>Respuesta del proceso de negocio</b>
<ol style="list-style-type: none"> <li>Solicita el análisis de la molécula</li> <li>Compara los resultados</li> </ol>	- Se realiza el análisis de la molécula
<b>Flujos alternativos</b>	-
<b>Acción del actor Bioinformatico</b>	<b>Respuesta del proceso de negocio</b>
<ol style="list-style-type: none"> <li>Recibe la solicitud</li> <li>Analiza la molécula</li> <li>Devuelve el resultado</li> </ol>	- Análisis de la molécula para la comparación posterior
<b>Mejoras propuestas</b>	-



La figura 3 realiza una representación del Modelo Conceptual.

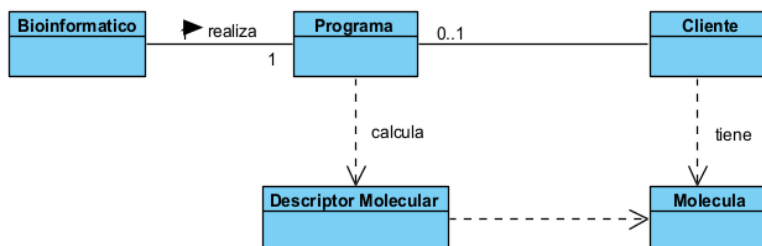


Figura 3. Modelo conceptual

**Bioinformático:** profesional que aplica técnicas de informática para analizar y entender datos biológicos. Su trabajo implica el manejo y análisis de grandes volúmenes de información genética, proteómica, y otras áreas de la biología molecular.

**Programa:** es un software diseñado para realizar tareas específicas en el campo de la bioinformática. En este contexto, el programa calcula los descriptores moleculares.

**Cliente:** usuario que necesita procesar moléculas.

**Molécula:** representaciones digitales de moléculas biológicas, como proteínas o nucleótidos.

**Descriptor molecular:** medidas cuantitativas que describen las propiedades físicas y químicas de una molécula. Estos pueden incluir aspectos como la polaridad, la hidrofobicidad, o la estabilidad térmica de la molécula.

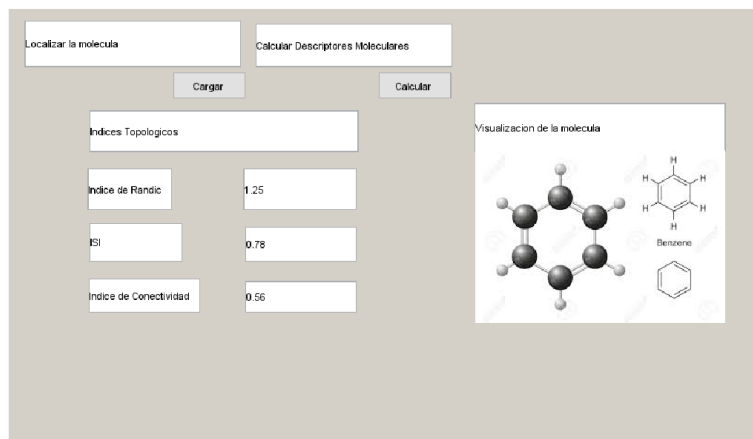


Figura 4. Prototipo de interfaz gráfica.

Para comprobar el correcto funcionamiento de los requerimientos definidos se diseñó una estrategia de prueba.



### **Método de Caja Blanca**

1. La interfaz de usuario fue diseñada de manera intuitiva para facilitar la interacción del cliente con el software, permitiéndole realizar solicitudes de análisis y comparar resultados de forma sencilla.
2. Se llevó a cabo un proceso de optimización en el software para garantizar un procesamiento eficiente de los datos moleculares, permitiendo realizar el análisis requerido en un tiempo razonable.
3. Se implementaron medidas de seguridad y privacidad en el almacenamiento de datos del cliente, asegurando un acceso seguro a los mismos y facilitando la comparación con resultados anteriores.
4. El software fue capaz de generar informes detallados sobre los análisis realizados, proporcionando al cliente la información necesaria para revisar y comparar los resultados con otros estudios previos.
5. Se realizaron pruebas exhaustivas para garantizar el correcto funcionamiento del software y la integridad de los resultados, corrigiendo cualquier error o fallo encontrado durante la ejecución de los casos de prueba.

### **Método de la Caja Negra**

1. Se identificaron las funciones principales del software, que incluían la recepción de solicitudes de análisis de moléculas por parte del cliente, el procesamiento y análisis de las moléculas por parte del bioinformático, la entrega de resultados al cliente y la comparación de resultados con estudios previos.
2. Se diseñaron casos de prueba que cubrían diferentes escenarios posibles, como el caso de prueba 1 donde el cliente enviaba una solicitud válida y el bioinformático procesaba la molécula correctamente, el caso de prueba 2 donde el cliente enviaba una solicitud inválida y el software manejaba el error adecuadamente, el caso de prueba 3 donde el cliente comparaba los resultados correctamente con estudios previos, y el caso de prueba 4 donde el cliente realizaba una solicitud con datos incompletos y el software solicitaba la información faltante.
3. Se ejecutaron los casos de prueba y se verificó que el software cumplía con las funcionalidades esperadas en cada escenario.
4. Se analizaron los resultados de las pruebas y se corrigieron cualquier error o fallo encontrado durante la ejecución.
5. Se repitió el proceso de prueba con diferentes datos de entrada para asegurar que el software funcionaba correctamente en diferentes situaciones.

La tabla 2 muestra una representación del diseño de pruebas funcionales.



**Tabla 2.** Diseño de pruebas funcionales

Escenario	Descripción	Variable	Respuesta del Sistema	Flujo central
Análisis de moléculas	El cliente solicita un análisis de una molécula. El bioinformático realiza el análisis y entrega los resultados. El cliente compara estos resultados con estudios previos.	Solicitud de análisis	Recepción de solicitud de análisis por parte del cliente.	Cliente solicita análisis de molécula.
		Realización del análisis	Procesamiento y análisis de la molécula por parte del bioinformático.	Bioinformático analiza la molécula.
		Entrega de resultados	Presentación de los resultados del análisis al cliente.	Bioinformático entrega resultados al cliente.
		Comparación de resultados	Cliente compara los resultados obtenidos con estudios previos para evaluar similitudes o diferencias.	Cliente compara resultados con estudios anteriores.

La tabla 3 muestra una Descripción de Variables.

**Tabla 3.** Descripción de Variables.

No.	Nombre de Campo	Clasificación	Valor Nulo	Descripción
1	Molécula	Campo de texto	no	Acepta solo archivos en formato .fasta
2	Resultados del Análisis	Campo numérico	no	Acepta números para los resultados de los análisis





## Conclusiones

Siguiendo el método de la Estrella de Bhoemy Turner, según los resultados obtenidos se puede apreciar que una de las mejores metodologías para trabajar es la XP, debido a su agilidad y adaptabilidad a pequeños entornos.

Mediante el modelo de negocio planteado se pueden apreciar las acciones que realizan los diferentes actores, así como la respuesta del programa.

El software ha sido diseñado para procesar solicitudes de análisis de moléculas, calcular descriptores moleculares, devolver resultados y cumplir con requisitos no funcionales como soportar múltiples solicitudes simultáneas y tener un tiempo de respuesta máximo de un minuto. Se ha logrado una interacción efectiva entre el cliente y el bioinformático, donde el cliente puede solicitar análisis, el bioinformático realiza el análisis y entrega resultados para su comparación con estudios previos.

El software ha sido optimizado para un procesamiento eficiente de datos moleculares y se han implementado medidas de seguridad y privacidad en el almacenamiento de datos del cliente, garantizando un acceso seguro y confiable.

Se han realizado pruebas exhaustivas, tanto de caja blanca como de caja negra, para garantizar el correcto funcionamiento del software, corrigiendo errores y fallos encontrados durante la ejecución de los casos de prueba.

La creación de este software ha sido guiada por una planificación detallada, implementación cuidadosa de requisitos funcionales y no funcionales, interacción efectiva entre los actores involucrados, optimización del procesamiento de datos, seguridad en el almacenamiento, generación de informes detallados y pruebas rigurosas para garantizar la calidad y fiabilidad del software en el análisis de moléculas.

## Conflictos de intereses

Los autores no poseen conflictos de intereses.

## Contribución de los autores

1. Conceptualización: Claudia M. Cardoso Gallardo, María Teresa de la Cruz Padilla
2. Curación de datos: Claudia M. Cardoso Gallardo
3. Análisis formal: María Teresa de la Cruz Padilla
4. Investigación: Claudia M. Cardoso Gallardo
5. Metodología: Claudia M. Cardoso Gallardo
6. Software: Claudia M. Cardoso Gallardo



7. Supervisión: María Teresa de la Cruz Padilla
8. Validación: María Teresa de la Cruz Padilla
9. Visualización: María Teresa de la Cruz Padilla
10. Redacción – borrador original: Claudia M. Cardoso Gallardo, María Teresa de la Cruz Padilla
11. Redacción – revisión y edición: Claudia M. Cardoso Gallardo, María Teresa de la Cruz Padilla

## Financiamiento

La investigación no requirió fuente de financiamiento externa.

## Referencias

- Alves, V. M., Braga, R. C., Muratov, E. N., & Andrade, C. H. (2018). Quimioinformática: uma introdução. *Química Nova*, 41, 202-212. <https://www.scielo.br/j/qn/a/K6fzXSJWRGLmqzkfbfBsqrj/?format=pdf&lang=pt>
- Armijos-Jaramillo, V., Yeager, J., Muslin, C., & Perez-Castillo, Y. (2020). SARS-CoV-2, an evolutionary perspective of interaction with human ACE2 reveals undiscovered amino acids necessary for complex stability. *Evolutionary Applications*, 13(9), 2168-2178. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/eva.12980>
- Borroto, O. M. R., Díaz, Y. H., de la Vega, J. M. G., Abalo, R. d. C. G., & Ponce, Y. M. (2012). Comparación de modelos novedosos de proximidad en Quimioinformática. *Afinidad*, 69(560). <https://www.raco.cat/index.php/afinidad/article/download/268518/356110>
- Campos, S. G., & Martínez, L. F. F. (2015). Programación extrema: prácticas, aceptación y controversia. *Cultura científica y tecnológica*(15). <http://erevistas.uacj.mx/ojs/index.php/culcyt/article/download/512/491>
- Medina-Franco, J. L., Fernández-de Gortari, E., & Naveja, J. J. (2015). Avances en el diseño de fármacos asistido por computadora. *Educación química*, 26(3), 180-186. <https://www.sciencedirect.com/science/article/pii/S0187893X15000294>
- Torres, N., Landau, L., Bamonte, E., Di Giacomo, M., Erausquin, P., Fornaso, C., & Monteserin, H. (2005). Fuerzas intermoleculares y propiedades físicas de compuestos orgánicos: una estrategia didáctica. *Educ. quim.*, 16(extraordinario), 129-136. <https://revistas.unam.mx/index.php/req/article/download/66076/57988/192782>
- Trujillo, A. G. P. (2011). La quimioinformática, una herramienta eficiente para desarrollar los medicamentos del futuro. *Teoría y praxis investigativa*, 6(1), 77-86. <https://dialnet.unirioja.es/descarga/articulo/3726081.pdf>

