

Prototipo para la clasificación de Proyectos de Ciencia Ciudadana utilizando inteligencia artificial: Experiencias en la Universidad del Rosario

Malgorzata Lisowska¹, Blanco Castillo Humberto²

Palabras claves

Ciencia ciudadana, inteligencia artificial, procesamiento de lenguaje natural, ciencia abierta

Citizen science, artificial intelligence, Natural Language Processing, Technological Innovation

Eje temático

Inteligencia artificial (IA) aplicada a la Ciencia Abierta

Resumen

Este artículo examina un prototipo para evaluar y clasificar proyectos de ciencia ciudadana, integrando una rúbrica de evaluación y un modelo de Procesamiento de Lenguaje Natural (PLN) personalizado. El enfoque, basado en la inteligencia artificial (IA), subraya su relevancia en la clasificación y abre nuevas perspectivas para la investigación en ciencia abierta.

El análisis comienza revisando el estado global de la ciencia ciudadana y los avances tecnológicos que permiten la evaluación y procesamiento de estos proyectos. Se destaca especialmente la experiencia de la Universidad del Rosario en la implementación de su modelo de ciencia abierta, enfrentando desafíos en la clasificación de proyectos y utilizando la IA para desarrollar un prototipo que determina la pertinencia de los proyectos dentro de la ciencia ciudadana. Este prototipo utiliza una rúbrica complementada con preguntas dinámicas que facilitan una clasificación precisa.

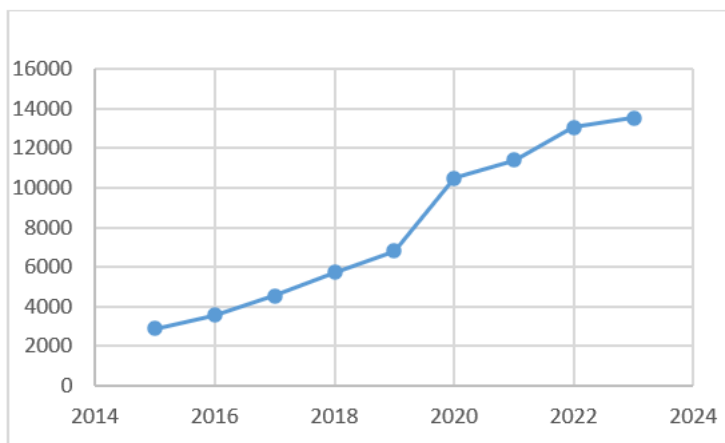
Se concluye discutiendo las ventajas de esta tecnología y su potencial para mejorar la evaluación de proyectos con soporte en inteligencia artificial.

Introducción

La ciencia ciudadana, caracterizada por la participación activa de no científicos en procesos de investigación científica, ha ganado relevancia como un enfoque inclusivo y democratizador para la generación de conocimiento. Además, ha experimentado un crecimiento notable en la última década (Tabla 1), lo cual plantea desafíos significativos para su evaluación y clasificación efectiva.

1 CRAI Universidad del Rosario - Colombia margarita.lisowska@urosario.edu.co

2 CRAI Universidad del Rosario - Colombia humberto.blanco@urosario.edu.co

Tabla 1 - Crecimiento de publicaciones de ciencia ciudadana

Por una parte, el desarrollo de plataformas de ciencia ciudadana ha potenciado la participación masiva de voluntarios en proyectos científicos y mejorado los métodos para la recopilación y clasificación de datos. Plataformas como Zooniverse, la mayor del mundo, con más de 2,000 proyectos y más de 1.6 millones de voluntarios en campos que van desde la astronomía hasta la zoología; iNaturalist, la cual permite mapear y compartir observaciones de biodiversidad, contabilizando más de 50 millones de observaciones hasta 2021; eBird, para registrar observaciones de aves, contribuyendo significativamente al modelado de distribuciones y patrones migratorios globales; Foldit, un juego en línea lanzado en 2008, que a través del juego ha logrado importantes avances en la biología de las proteínas. No solo demuestran el poder de la colaboración masiva y la tecnología para avanzar en la ciencia, sino que también subrayan cómo la participación pública puede transformar la investigación científica y aumentar el conocimiento colectivo de manera significativa y duradera.

Por otro lado, La inteligencia artificial (IA) está revolucionando la ciencia ciudadana al automatizar y mejorar la recopilación y clasificación de grandes volúmenes de datos, especialmente en proyectos relacionados con la identificación de especies y el monitoreo ambiental. Los algoritmos de reconocimiento de imágenes, por ejemplo, facilitan la identificación automática de especies o condiciones ambientales a partir de fotografías cargadas por ciudadanos, aumentando la eficiencia y reduciendo la carga de trabajo manual. Además, los algoritmos de aprendizaje automático están permitiendo el descubrimiento de patrones y correlaciones en los datos que serían difíciles de detectar por humanos, contribuyendo a nuevos descubrimientos científicos y mejorando la comprensión de fenómenos complejos.

Esta sinergia entre la participación humana y la inteligencia artificial no solo potencia la capacidad de campos de conocimiento donde los proyectos de ciencia ciudadana contribuyen a la ciencia formal, sino que también mejora la experiencia y aprendizaje de los participantes, como demuestran estudios de Ceccaroni, L. et al. (2019) y Rafner, J. et al. (2021). También, este potencial para introducir la ciencia ciudadana en más campos del conocimiento, genera un nuevo reto, haciendo más lento el proceso de clasificación y evaluación manual de este tipo de proyectos.

La Universidad del Rosario ha adoptado la ciencia abierta como parte esencial de su quehacer, enfocándose en una investigación de alta calidad que es inclusiva, colaborativa y transparente, destinada a impactar positivamente en la sociedad. Desde hace más de una década, ha participado activamente en el movimiento de acceso abierto, implementando herramientas clave como el repositorio E-docUR y el portal de revistas de acceso abierto, y fue pionera en Colombia al establecer una política institucional de acceso abierto en 2017. En 2020, la universidad intensificó sus esfuerzos creando un modelo de ciencia abierta orientado por ocho pilares fundamentales, enfocados en aspectos como la comunicación académica, los datos de investigación, las métricas de nueva generación, la integridad científica y la ciencia ciudadana. Además, ha desarrollado infraestructura y herramientas adicionales para facilitar la ciencia abierta, culminando en la creación de un Portal Institucional de Ciencia Abierta que refleja su compromiso continuo y promueve la iniciativa a nivel regional.

Luego de establecer el modelo se inició la realización de un diagnóstico en una unidad académica que sirviera como ejemplo para evidenciar la materialización de su modelo la ciencia abierta en la Universidad, con ello se identificaron iniciativas y alcances en los diferentes pilares del modelo incluyendo aquellas iniciativas de investigación que involucran la participación activa de la comunidad. En este proceso de diagnóstico se encontraron algunos desafíos al invitar a la comunidad académica a presentar sus proyectos de ciencia ciudadana, aun cuando se realizó una caracterización de la ciencia ciudadana en la UR.

Para abordar estos desafíos, se propuso un prototipo para automatizar y mejorar el proceso de clasificación y evaluación de proyectos de ciencia ciudadana mediante el uso de tecnologías de procesamiento de lenguaje natural, junto con una rúbrica de evaluación. Su objetivo es proporcionar una herramienta eficaz que permita evaluar rápidamente si un proyecto cabe dentro de la categoría de ciencia ciudadana basado en una serie de criterios preestablecidos en la caracterización de este eje dentro del modelo de ciencia abierta.

El prototipo propuesto

Aprovechando las ventajas actuales de los modelos de aprendizaje para analizar y comprender el contenido de las propuestas de proyectos de ciencia ciudadana y, aprovechando además su capacidad para identificar objetivos clave, metodologías propuestas, y recursos necesarios descritos en los documentos de proyecto, se incorporó un sistema de clasificación basado en una rúbrica predefinida que incluye criterios como innovación, impacto potencial, viabilidad técnica, y alineación con los objetivos de la ciencia ciudadana definidos en el modelo de ciencia abierta de la Universidad del Rosario; además de la clasificación, ofrece retroalimentación y sugerencias de mejora basadas en el análisis realizado, para ayudar en el perfeccionamiento de las propuestas.

Para ello, utilizando ChatGTP se personalizó GTP (RubrikCivitas), con un prompt en el que se define la rúbrica de evaluación basado en 5 criterios principales a los cuales se les otorgó un puntaje:

- Diálogo de saberes entre academia y comunidades (20 puntos)
- Investigación trans e interdisciplinar (20 puntos)
- Fortalecimiento del trabajo investigativo (20 puntos)
- Involucramiento de múltiples fuentes de información (20 puntos)
- Promoción de la ciudadanía activa y la democratización del saber (20 puntos)

Además, se configuró el prompt para que además del resultado de evaluación, el GTP indicara:

- Identificación de las fortalezas y debilidades del proyecto en relación con cada criterio de la rúbrica.
- Ejemplos específicos del proyecto que respalden las evaluaciones del GPT-4.
- Recomendaciones para mejorar el proyecto.

Los resultados preliminares indican una mejora significativa en la eficiencia y calidad de la evaluación de proyectos. Se espera que futuras iteraciones del prototipo integren capacidades de aprendizaje automático más avanzadas y que se expanda su uso a otras facultades y Escuelas.

El prototipo implementado se encuentra en fase de exploración de manera privada, el objetivo final es automatizar el proceso, adicionando otras herramientas de tal manera que pueda ser accesible para cualquier usuario.

Resultados

Para realizar las validaciones se revisaron 17 proyectos (Tabla 2), 12 de los cuales se habían identificado previamente como proyectos de ciencia ciudadana, se contó con un 52% de efectividad respecto a la clasificación de dichos proyectos, en todos los casos por no contar con información suficiente para determinar si los proyectos cumplen o no con los aspectos indicados en la rúbrica de evaluación.

Tabla 2 - Resultados identificación de proyectos de ciencia ciudadana

Número	Proyecto	¿Es proyecto de ciencia ciudadana?
1	El impacto de estrategias de integración asistencial sobre redes integradas de servicios de salud en distintos sistemas de salud de América Latina	Sí
2	Herramientas para la construcción de paz y convivencia en contextos escolares	Sí
3	Educación inicial saludable inclusiva y diversa en el sector El Codito en los cerros orientales de la UPZ 9 (Verbenal) localidad de Usaquén	Sí
4	Modelo de formación en sexualidad y derechos sexuales y reproductivos para personas con discapacidad intelectual	Sí
5	La Colombia imaginada trazos de paz: la literatura infantil como experiencia pedagógica en educación superior	Sí
6	Relación entre contaminación del aire por material particulado (PM10) y morbilidad respiratoria en niños menores de 5 años en la localidad de Kennedy	No
7	Desarrollo de capacidades en jóvenes con discapacidad cognitiva para la gestión del riesgo	No
8	Estrategias de manejo y gestión del riesgo en la localidad de San Cristóbal Sur	No
9	Fortalecimiento de la red de prestadores de servicios de salud en la localidad de Bosa	No
10	Impacto de los espacios de atención ciudadana en la salud mental comunitaria	No

11	Gestión integral de la calidad del agua en la localidad de Chapinero	No
12	Promoción de hábitos saludables en la población escolar de la localidad de Fontibón	No
13	Exploring precipitation toxoplasmosis in Colombia	Si
14	Cuidadores final	No
15	atlas.admon.latam	No
16	Informality	No
17	Replication Data for Network Topology in Decentralized Finance	No

Conclusiones y trabajo futuro

El desarrollo y la implementación del prototipo basado en ChatGPT en la Universidad del Rosario han marcado un avance significativo en la efectividad con la que se pueden evaluar y clasificar los proyectos de ciencia ciudadana, al facilitar un proceso más rápido y menos subjetivo. Al proporcionar evaluaciones detalladas y sugerencias constructivas basadas en una rúbrica estandarizada, el sistema puede incentivar a los proponen-tes a ajustar y mejorar sus propuestas en fases preliminares de la presentación del proyecto.

Este tipo de iniciativas, se pueden incorporar a algoritmos de aprendizaje profundo de código abierto, que puedan mejorar la capacidad del sistema para entender y procesar lenguaje natural, permitiendo una evaluación aún más precisa de las propuestas.

La evaluación y clasificación de propuestas de ciencia ciudadana, es apenas una muestra del potencial de este tipo de modelos para explorar su efectividad en la evaluación en otros contextos de la ciencia abierta.

Se debe automatizar el proceso y abrirlo a la comunidad a través de la creación de interfaces de usuario adaptadas a las necesidades específicas de diferentes grupos de evaluadores y proponentes, mejorando la accesibilidad y usabilidad del sistema.

En conclusión, la implementación del prototipo basado en ChatGPT en la Universidad del Rosario subraya las ventajas significativas de integrar tecnologías avanzadas en la evaluación de proyectos de ciencia ciudadana. Este enfoque no solo mejora la precisión y la eficiencia en la evaluación de proyectos basados en rúbricas, sino que también demuestra el potencial de la inteligencia artificial para transformar procesos tradicionalmente manuales en sistemas automatizados y altamente eficaces. Según Ceccaroni et al. (2019) y Rafner et al. (2021), el uso de algoritmos avanzados permite una interpretación más profunda y una evaluación más objetiva de los criterios establecidos en las rúbricas, facilitando así decisiones más informadas y justas. Este prototipo representa un paso adelante hacia la optimización de la evaluación de proyectos, ofreciendo un modelo replicable y escalable que puede ser adaptado a diferentes campos y disciplinas científicas, promoviendo una mayor inclusión y democratización en la generación de conocimiento científico.

Bibliografía

Portal institucional de Ciencia Abierta de la Universidad del Rosario. Accedido 29 de abril de 2024. <https://cienciaabierta.urosario.edu.co/>

Ceccaroni, L., Bibby, J., Roger, E., Flemons, P., Michael, K., Fagan, L., & Oliver, J. (2019). Opportunities and Risks for Citizen Science in the Age of Artificial Intelligence. *Citizen Science: Theory and Practice*. <https://doi.org/10.5334/cstp.241>

Rafner, J., Gajdacz, M., Kragh, G., Hjorth, A., Gander, A., Palfi, B., Berditchevskaia, A., Grey, F., Gal, Y., Segal, A., Walmsley, M., Miller, J., Dellerman, D., Haklay, M., Michelucci, P., & Sherson, J. (2021). Revisiting Citizen Science Through the Lens of Hybrid Intelligence. *ArXiv*, abs/2104.14961.

Malgorzata Lisowska, directora del Centro de Recursos para el Aprendizaje y la Investigación – CRAI, de la Universidad del Rosario en Bogotá.

Magister en Bibliotecología e Información Científica, Universidad Jagiellona de Cracovia, Polonia. Especialista en Administración de Empresas, Universidad del Rosario. Especialista en Gerencia y Gestión Cultural, Universidad del Rosario. Amplia experiencia en bibliotecas públicas y universitarias, con énfasis en gestión y evaluación bibliotecaria y en implementación de nuevas tecnologías. investigadora en el proyecto de la Creación De La Biblioteca Digital Colombiana BDCOL, coordinación de proyectos internacionales como CoLaBoRa (Comunidad Latinoamericana de Bibliotecas y Repositorios Digitales) y en “LA Referencia” patrocinado por de la RedClara y el BID.

Humberto Blanco Castillo, jefe de Innovación y Proyectos del Centro de Recursos para el Aprendizaje y la Investigación – CRAI, de la Universidad del Rosario en Bogotá.

Ingeniero de sistemas, especialista en gerencia de proyectos TIC. Experto en el desarrollo de proyectos enfocados a la implementación, visibilidad e interoperabilidad de repositorios institucionales, así como el desarrollo de soluciones basadas en software libre para la gestión de bibliotecas. Actualmente lidera las estrategias para promover la visibilidad de la producción institucional en acceso abierto, la generación iniciativas y gestión proyectos de base tecnológica que apoyan a los procesos de innovación del CRAI.