Published on *Hoy en el TEC* (https://www.tec.ac.cr/hoyeneltec)

Según los investigadores, cada vez se está más cerca de costruir un sistema de inteligencia artificial capaz de identificar todas las especies de plantas del mundo. Estas fotografías son un ejemplo de las imágenes que BioMachina utilizó para construir su modelo de Machine Learning. Imagen cortesía de R. González.

Machine Learning

Equipo BioMachina es cuarto mundial al superar adversidades en sistema de inteligencia artificial

15 de Septiembre 2022 Por: Johan Umaña Venegas [1]

Grupo de Maestría en Computación desarrolló eficiente modelo que usa menor poder computacional y da resultados competentes

La conferencia PlantCLEF 2022, sobre identificación de plantas con computadoras, se desarrolló en setiembre, en Bolonia, Italia

Rubén González y José Mario Carranza, en el evento CLEF 2022, donde fueron a presentar los resultados de su participación en la competencia PlanCLEF 2022. Imagen cortesía.

Al enfrentarse a las adversidades para el **acceso a equipo de computo de mayor capacidad –en comparación con los recursos en países desarrollados–** el equipo **BioMachina desarrolló un enfoque novedoso y eficiente**, que no solo llama la atención de expertos internacionales, sino que además sirvió para quedar **cuartos en el ranking mundial de la competencia de identificación automática de plantas PlantCLEF 2022 [2].**

La investigación fue presentada en el artículo titulado "*Extreme automatic plant identification under constrained resources* [3]" (**Identificación automática extrema de plantas con recursos limitados**), con la autoría del doctor José Mario Carranza Rojas, docente de la Maestría en Computación con énfasis en Ciencias de la Computación [4] del TEC, y cinco estudiantes de ese posgrado: Rubén González Villanueva, Kelvin Jiménez Morales, Kevin Quesada Montero, Esteban Esquivel Barboza y Nicole Carvajal Barboza.

> “Este artículo es clave para democratizar el acceso a investigadores que quieran participar en esta área de investigación, ya que no todos cuentan con los caros equipos que paulatinamente se van ocupando, dado que ahora sí estamos en vías de identificar cualquier especie de plantas en el mundo”, comenta Carranza.

El investigador se refiere a que cada vez **se está más cerca de crear un sistema de inteligencia artificial capaz de reconocer todas las especies de plantas conocidas en el mundo**, solo con el uso de imágenes. En esta edición de competencia de **PlantCLEF se utilizó un set de 80.000 especies, y en el mundo hay unas 400.000 especies.**

Sin embargo, estos avances tienen sus desafíos y es que los modelos cada vez son más

complejos y requieren de mayor poder computacional. Limitando, como muchas veces sucede, la participación de científicos de diferentes partes del mundo en el desarrollo de técnicas de avanzada.

“Actualmente existen modelos de Machine Learning muy poderosos, como CNN y Transformers, que necesitan un alto nivel computacional para poder entrenarlos y esto es muy caro.
El principal aprendizaje de nuestra investigación fue inventar una arquitectura con la cual se puede reducir el tamaño de los modelos en más de cinco veces, lo cual hace más factible entrenarlos ya que requiere menos poder computacional, y obtener resultados muy similares”, agrega Rubén González.

[5]

Estudiante gana reto mundial de inteligencia artificial y deja en alto nombre del TEC [5]

Ingenio para sobrellevar los retos

En un inicio BioMachina buscó desarrollar un modelo más complicado, pero se topó con la realidad de que en el TEC no contaban con los clústers de ordenadores necesarios para cargar tantas imágenes y correr los gigantescos modelos de Machine Learning que consideraban necesario para la competencia. También, que las facilidades del Centro Nacional de Alta Tecnología (CeNAT) ya están sobrecargadas.

La respuesta sorprendió a investigadores internacionales, pues desarrollaron un modelo más liviano y rápido, que no solo cumplió a cabalidad con los requerimientos y rindió a nivel de ser el cuarto lugar a nivel mundial, sino que además corrió mucho más rápido de lo esperado.

“Los resultados son muy satisfactorios: por ejemplo, un modelo que típicamente tendría unos 160 millones de parámetros, con nuestro esquema tiene 28 millones, permitiéndonos entrenar con todas las especies e imágenes, a pesar de no contar con el equipo óptimo, en solo una fracción del tiempo, y obteniendo los mismos resultados que si hubiésemos corrido el mismo modelo sin nuestra capa”, ahonda Carranza.

Esa efectividad en el uso de los recursos y el desarrollo de una técnica nueva, a la que llamaron “*learned hierarchy*”, o jerarquía aprendida,

fue lo que maravilló a los organizadores del evento. El modelo reduce la cantidad de parámetros al inferir una jerarquía similar a la taxonomía de especies de plantas conocidas (como géneros, familias, etc.), pero basada solo en características visuales.

“No fuimos el primer lugar en la competencia, pero en la conferencia se destacó nuestra participación, por el impacto y el potencial que tiene en el futuro de la Inteligencia Artificial nuestra innovación. Esto se puede ver en las conclusiones de la competencia, donde dice que la taxonomía puede aprender mucho de la jerarquía aprendida, ya que es ‘prometedora y eficiente’”, contó González.

BioMachina espera seguir mejorando la técnica desarrollada para obtener mejores resultados, y además seguir participando en esta y otras competencias similares.

Source URL (modified on 10/18/2022 - 16:13): https://www.tec.ac.cr/hoyeneltec/node/4289

Enlaces
[1] https://www.tec.ac.cr/hoyeneltec/users/johan-umana-venegas
[2] https://www.imageclef.org/PlantCLEF2022
[3] https://scholar.google.com/citations?view_op=view_citation&hl=en&user=P_u5iNcAAAAJ&sortby=pubda
[4] https://www.tec.ac.cr/programas-academicos/maestria-computacion
[5] https://www.tec.ac.cr/hoyeneltec/2020/09/04/estudiante-gana-reto-mundial-inteligencia-artificial-deja-alto-nombre-tec