

G

Grupo PaRMa-TEC promueve investigación en aprendizaje automático

*Saúl Calderón-Ramírez
 **Esteban Arias-Méndez
 ***Mauricio Arroyo-Herrera



Investigadores y colaboradores del grupo PaRMa con algunos invitados especiales en la conferencia internacional IWObI 2018, en el Campus del TEC en San Carlos. La actividad fue co-organizada por PaRMa. Fuente: Grupo PaRMa.

Aprendizaje automático

El aprendizaje automático (*machine learning*) se define como la subdisciplina dentro de la inteligencia artificial que estudia algoritmos y modelos matemáticos que aprenden a partir de un conjunto de datos, para clasificar o estimar nuevos datos y encontrar patrones que apoyen o automaticen la toma de decisiones en distintos ámbitos. El aprendizaje automático es una herramienta cada vez más popular en distintas áreas científicas, desde la biología hasta la sociología, y con un número creciente de aplicaciones en medicina, economía, industria y en los negocios. La creciente popularidad del aprendizaje automático es catalizada por las siguientes razones:

1. La creciente disponibilidad de datos digitales, generada por una cada vez mayor cantidad de dispositivos y medios de captura de datos: redes sociales, sensores equipados en celulares, cámaras, etc.
2. Una caída en el costo de almacenamiento y transmisión de los datos, con redes inalámbricas y cableadas más rápidas y medios de almacenamiento masivos y a gran escala.
3. Aumento en la capacidad de cómputo y descenso de sus costos: los procesadores modernos incorporan gran cantidad de unidades de procesamiento, capaces de procesar en paralelo grandes cantidades de datos.

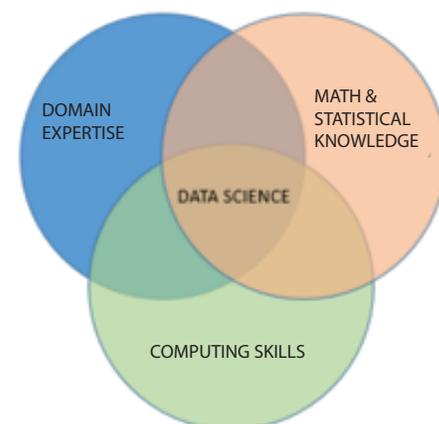
En áreas con mayor facilidad y disponibilidad para generar y almacenar datos, como la economía, el aprendizaje automático ha sido extensamente utilizado en modelos de predicción de comportamientos micro y macroeconómicos, viabilidad financiera, etc. En la ingeniería, el diseño de dispositivos de distintos tipos es, cada vez más, asistido por sistemas basados en el aprendizaje automático para calibrar parámetros y distintas variables de diseño de

forma inteligente. Además, incorpora funcionalidades adaptativas y completamente automáticas para los usuarios finales. En la medicina, el aprendizaje automático empieza a usarse de forma extensiva para crear modelos que permitan asistir en el diagnóstico y formulación de tratamientos personalizados a los pacientes, lo cual se conoce recientemente como medicina de precisión. De forma similar en la agricultura de precisión, el aprendizaje automático se populariza como herramienta para entender de forma más rápida y precisa el estado de un cultivo y asistir en la toma de decisiones en cuanto a la fertilización y estimación de las cosechas.

La aplicación del aprendizaje automático a distintas áreas abre la necesidad de agilizar la interacción entre los distintos dominios de aplicación (dominio específico), las ciencias de la computación y la matemática, además de ingenierías relacionadas en el diseño e implementación de soluciones basadas en el aprendizaje automático. El diagrama de la Figura 1, muestra cómo esta interacción entre la matemática o estadística, las ciencias de la computación y el dominio específico generan lo que recientemente se conoce como ciencias de los datos.

Grupo PaRMa

El grupo PaRMa (Pattern Recognition and Machine Learning) es un grupo multidisciplinario conformado principalmente por profesores investigadores del Instituto Tecnológico de Costa Rica (TEC), en el contexto del programa eScience. Fue conformado en agosto del 2016 por los profesores Saúl Calderón y José Mario Carranza y nació como parte de la inquietud de la Escuela de Computación de contar con una colaboración más integral en iniciativas de investigación y desarrollo en las áreas de reconocimiento de patrones (*pattern recognition*) y aprendizaje automático (*machine learning*). La misión del grupo PaRMa sigue los tres ejes del TEC:



library.osu.edu/blogs/it/where-does-data-science-fit-in/

1. **Docencia:** Generar espacios de discusión y propuestas para transformar los programas académicos de las carreras relacionadas en el desarrollo y aplicación del aprendizaje automático. Esto lo hemos podido alcanzar mediante la organización de escuelas cortas, seminarios, simposios, charlas y otras actividades que permitan formar a estudiantes universitarios y público interesado en los temas de interés del grupo.
2. **Investigación:** El grupo PaRMa apoya la creación, propuesta y puesta en marcha de proyectos de investigación relacionados con el desarrollo y aplicación de técnicas de aprendizaje automático. Además, genera redes de colaboración con universidades de renombre internacional, instituciones públicas y privadas, necesarias en la búsqueda de recursos para tales iniciativas. Gracias al interés y apoyo de los miembros se han generado y publicado un número considerable de publicaciones.
3. **Extensión:** La aplicación y transferencia de resultados de proyectos realizados en el con-

texto del grupo PaRMA es un objetivo importante, como paso final en la generación de valor para la sociedad costarricense. La transferencia de conocimiento mediante múltiples actividades académicas es el más claro ejemplo alcanzado por el grupo.

Líneas de investigación

Las siguientes son las tres líneas de investigación principales del grupo.

- **Agricultura de precisión.** El uso de datos de imágenes térmicas e infrarrojas tomadas de drones para estimar la cobertura vegetal y las propiedades químicas del suelo, es un ejemplo del uso de datos para la toma de decisiones en la agricultura, con investigación desarrollada dentro del grupo PaRMA.
- **Medicina de precisión.** La recopilación de datos de imágenes de diferentes fuentes (rayos X, resonancia magnética, tomografía computarizada, etc.), datos fisiológicos (EEG, ECG) y también datos de comportamiento, permite una toma de decisiones más precisa y eficiente para médicos, radiólogos, patólogos y ciencias de la vida. Ejemplos de investigaciones desarrolladas en el grupo PaRMA son la estimación de la edad ósea mediante el uso de imágenes de rayos X y la segmentación y seguimiento de las células mediante videos de microscopía basada en fluorescencia, para evaluar el impacto de la quimioterapia en el tejido canceroso.
- **Conservación de la biodiversidad.** A medida que el cambio climático se convierte en un desafío a corto plazo para la vida en nuestro planeta, las soluciones efectivas para el análisis de datos que alimentan la toma de decisiones

en la conservación de la biodiversidad también se vuelven más importantes. La identificación de plantas con cámaras regulares y macroscópicas con aprendizaje profundo es un ejemplo de investigación realizada desde PaRMA.

- **Reflexión del impacto de la inteligencia artificial (IA) y el aprendizaje automático.** La inteligencia artificial ha venido consolidándose a nivel mundial como herramienta a ser utilizada cada vez más en nuestra sociedad. Ello ha hecho que algunos expertos de la disciplina la consideren como el fundamento que moverá el quehacer de nuestra sociedad. Es así que encontramos IA en otras tendencias tecnológicas disruptivas de la transformación digital como: dispositivos autónomos, analítica aumentada, gemelos digitales, experiencias inmersivas, “smart spaces” y posiblemente otras más. Existe una concepción generalizada de que avances en robótica y *bots* digitales, reemplazarán gradualmente a los humanos. Por ejemplo, se considera que los vehículos autónomos reemplazarán a los taxistas y choferes, o que los drones con inteligencia artificial impactarán los procesos de logística o incluso podrán ser utilizados como armamento en la milicia. Para PaRMA, la dimensión social del impacto de la técnica que estudia no debe ser desatendida; y por tanto, desarrolla líneas de pensamiento sobre la naturaleza humana, social y política de la disciplina.

Además, el grupo PaRMA tiene por objetivo posicionarse a nivel nacional como un actor referente en temas de inteligencia artificial, aprendizaje automático y ciencias de los datos. Este impacto se ha visto hasta ahora de forma positiva, al haber sido citado en entrevistas con periódicos como La Nación y el Semanario Universidad, programas de radio como Amelia Rueda y programas de televisión como 7 días.

Comunidad de interés

Uno de los aspectos a destacar con respecto a la organización de PaRMA, es que ha venido usando el modelo de “comunidad de interés”. Este concepto ha surgido desde la Escuela de Computación y ha sido apoyado por las autoridades en ejercicio. Una comunidad de interés tiene como objetivo impulsar el desarrollo de actividades extracurriculares y co-curriculares en los estudiantes y docentes del TEC, gestionadas desde las escuelas a partir de organizaciones autónomas y autogestionadas basadas en la libertad de cátedra, que se constituyen por auto-determinación y que son producto de la socialización espontánea debido a intereses compartidos multidisciplinares.

Dado lo anterior, PaRMA se ha establecido como espacio auténtico y significativo para el ejercicio de

habilidades socio-emocionales, técnicas-científicas, de vinculación, de innovación, de investigación, de extensión y acción social; para ello, utiliza como motivador la creatividad que surge desde lo interno de la comunidad a partir de los intereses comunes de sus integrantes.

Las actividades ejecutadas desde PaRMA han contado con el apoyo de las autoridades en cada momento, para cada actividad desarrollada, tales como: la Vicerrectoría de Docencia, la Vicerrectoría de Investigación y Extensión y el Consejo de Escuela de Computación, entre otros. Ello en vista de que sus integrantes utilizan los mecanismos actuales disponibles en la institucionalidad, para favorecer la acción colectiva del grupo.

Actividades realizadas

Las siguientes son algunas de las actividades y resultados logrados desde la conformación del grupo PaRMA.

1. Docencia

a. **Programa de especialidad de ciencias de los datos:** es un programa de un año, dirigido a ingenieros y profesionales afines interesados en aprender los principios de las ciencias de los datos y transferir su aplicación, que ha dado inicio este 2019.

b. **Cursos en programas académicos del TEC:** los siguientes cursos han sido formulados y desarrollados en el contexto del grupo PaRMA, como electivas y ya aprobados por las instancias académicas.

- Introducción al reconocimiento de patrones.
 - Deep Learning*.
 - Introducción al procesamiento digital de imágenes.
 - Física computacional.
- c. Actividades cortas: las siguientes actividades cortas han sido organizadas o co-organizadas por el grupo PaRMA, de forma gratuita para la Comunidad TEC.
- PaRMA Flash-Talks*.
 - Simposio Internacional en Aplicaciones del Aprendizaje Automático, ediciones en 2017 y 2018.
 - Escuela de Veranillo en *Machine Learning*, 2018.
 - Ciclo de charlas sobre procesamiento de imágenes médicas, 2017.
 - Simposio en inteligencia artificial aplicada a imágenes médicas, 2019.
 - IEEE International Work Conference on Bioinspired Intelligence (IWOBI), 2018.
 - CARLA-BIP, 2019.

2. Investigación

Los siguientes proyectos de investigación inscritos en la VIE, están siendo ejecutados con apoyo del grupo.



Simposio de inteligencia artificial en imágenes médicas. Saúl Calderón, con el Dr. David Elizondo (derecha), de la Universidad De Montfort. Fuente: Grupo PaRMA.

- Clasificación y comparación de imágenes biomoleculares y médicas.
- Estudio de viabilidad de técnicas automáticas de visualización para ser usadas en la evaluación de impulsores en bombas de sangre.
- Análisis funcional genómico de células cancerosas por RNA de interferencia para la identificación de redes de regulación asociadas a proliferación y muerte en respuesta a quimioterapia genotóxica.
- Identificación de plantas de Costa Rica usando rasgos morfológicos y *Deep Learning*.
- Identificación de árboles de Costa Rica usando rasgos anatómicos de la madera y *Deep Learning*.

Además, más de una decena de estudiantes practicantes han hecho pasantías, gracias a las redes de colaboración en investigación construidas por el grupo, en universidades y centros de investigación extranjeros, tales como:

- Universidad de Toronto, Canadá.
- CIRAD, Francia.
- Universidad de Rovira i Virgili, España.
- Universidad Nacional del Litoral, Argentina.
- Universidad de la Coruña, España.

3. Extensión

- Formulación del curso corto “Introducción a la inteligencia artificial” para el Ministerio de Ciencia, Tecnología y Telecomunicaciones (MICITT), a impartirse en escuelas y colegios.
- Actualmente miembros del grupo PaRMA se encuentran en conversaciones con médicos del Hospital Nacional de Niños y el Hospital México para iniciar la prueba de los sistemas desarrollados en el grupo en tales hospitales.
- Talleres de Introducción a la programación a cientos de estudiantes de escuela y colegio mediante actividades de voluntariado desarrolladas por profesores y estudiantes del TEC de forma gratuita, con plataformas intuitivas como Arduino, RaspberryPi y Python.

Planes futuros

Los miembros del grupo PaRMA continúan colaborando y proponiendo proyectos de investigación y extensión, participando en foros sobre industria 4.0 y buscando alianzas con más colaboradores para el desarrollo de los temas de interés del grupo. Recientemente inició el Programa de Ciencia de los Datos, gestado desde PaRMA, con una duración de un año, para el cual se planeó iniciar con un grupo de al menos 20 personas. La demanda superó las expectativas y en este momento se están organizando al menos cinco grupos iniciales para atender a unas 100 personas. Actualmente se está trabajando con la organización de la conferencia internacional CARLA 2019, que

será en Costa Rica, donde el grupo tendrá una participación activa en la sesión especial denominada BIP (*BioInspired Processing*). ■

El grupo

M.Sc. Saúl Calderón Ramírez, Computación, coordinador.

Investigadores activos

M.Sc. (*in fieri*) Esteban Arias Méndez, Computación

Dr. Juan Luis Crespo Mariño, Mecatrónica

Dr. Juan Esquivel Rodríguez, Computación

Máster Luis Alexander Calvo Valverde, Computación

Dr. Martín Solís Salazar, Administración de Empresas

M.Sc. Felipe Meza Obando, Mecatrónica

M.Sc. Fabián Fallas Moya, Computación

Máster Efrén Jiménez Delgado, Computación

M.Sc. (*in fieri*) Esteban Pérez Hidalgo, Física

Contacto

<https://www.tec.ac.cr/grupo-investigacion/parma>

***Saúl Calderón Ramírez** es profesor e investigador de la Escuela de Computación del Instituto Tecnológico de Costa Rica (TEC). Magister Scientae en ingeniería eléctrica con énfasis en sistemas digitales de la Universidad de Costa Rica (UCR) y bachiller en computación e informática de la misma universidad. Entre sus especialidades están el procesamiento digital de señales (sonido, imágenes, video), reconocimiento de patrones, aprendizaje automático y computación paralela. Es el coordinador del PATTERN Recognition and Machine Learning Group (PaRMA-Group) y participante en iniciativas de análisis automático de imágenes biomédicas. Cuenta con once publicaciones en conferencias internacionales, en el área de procesamiento digital de señales, reconocimiento de patrones y aprendizaje profundo.

<https://orcid.org/0000-0001-9993-4388>

****Esteban Arias Méndez** es profesor, investigador y extensionista de la Escuela de Computación del TEC. Ingeniero en computación, forma parte del grupo de investigación PaRMA y coordina el grupo estudiantil *Singular*, grupo de investigación y extensión que gestiona el proyecto *Programación: la nueva alfabetización*. Adicionalmente trabaja en temas relacionados con inteligencia artificial y bioinformática sobre algoritmos alternos de comparación de rutas metabólicas. También es instructor certificado LPIC-1 de Linux. Desde 2014 utiliza la plataforma Arduino en sus cursos y ha brindado cursos de programación con Arduino abiertos a la comunidad de manera voluntaria.

<https://orcid.org/0000-0002-5600-8381>

*****Mauricio Arroyo Herrera** es profesor e investigador de la Escuela de Computación del TEC. Ha sido docente de la Universidad Estatal a Distancia (UNED). Ha colaborado como gestor curricular de la carrera de ingeniería en computación, fue coordinador de la Unidad Desconcentrada de la carrera de ingeniería en computación en el Centro Académico de San José y ha sido director de la Escuela de Ingeniería en Computación del TEC, que tiene a cargo siete unidades académicas. Ha participado en proyectos de investigación oficiales del TEC en temas de bases de datos genealógicas y sostenibilidad de las ciudades. Actualmente es candidato a doctor del programa de Gobierno y Políticas Públicas que imparte la Universidad de Costa Rica, donde desarrolla su tesis doctoral en el tema de ciudades sostenibles. Ha sido director de proyectos de tecnología de información en organizaciones del sector privado y público.

<https://orcid.org/0000-0001-6632-4831>



Escuela de Veranillo en *Machine Learning*, Centro Nacional de Alta Tecnología (CeNAT), Costa Rica. Fuente: Grupo PaRMA.



Alumni Day en la Facultad de Medicina de la Universidad de Toronto, Canadá. El coordinador, Saúl Calderón, expone las líneas de investigación del grupo PaRMA. Fuente: Grupo PaRMA.