

V

aniversario del grupo PaRMA/TEC Cinco años de avances en reconocimiento de patrones y aprendizaje automático

Esteban Arias-Méndez*
Coordinador Grupo PaRMA
esteban.arias@itcr.ac.cr



Ilustración 1. Resultado preliminar de rejuvenecer hojas de plantas de herbario automáticamente. Trabajo en proceso por Esteban Esquivel Barboza.

El grupo PaRMA (*Pattern Recognition and Machine Learning*) es un grupo multidisciplinario conformado principalmente por profesores investigadores del Instituto Tecnológico de Costa Rica (TEC). Fue conformado en agosto del 2016 por los profesores Saúl Calderón Ramírez y José Carranza, de la Escuela de Ingeniería en Computación. Luego se fueron sumando a la iniciativa los académicos Erick Mata, Luis-Alexánder Calvo-Valverde y Juan Esquivel Rodríguez, también de Computación; Juan Luis Crespo Mariño, de Mecatrónica; y Martín Solís, de Administración de Empresas, entre otros. Además se unieron varios estudiantes avanzados que se interesaron en aprender más sobre estos temas hasta llegar a ser asistentes de investigación.

Reconocimiento de patrones y aprendizaje automático

PaRMA nació ante la inquietud 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 ha seguido los tres ejes principales del TEC: docencia, investigación y extensión.

Producto de este trabajo, se han publicado más de 50 artículos científicos, asociados a miembros del grupo, se ha colaborado en la organización y realización de al menos tres conferencias académicas internacionales en Costa Rica (IWOB 2018, CARLA 2019 y BIP 2021), así como la asistencia y participación en múltiples conferencias y eventos internacionales. Se han organizado dos simposios de aprendizaje automático en el TEC, varias escuelas de verano gratuitas para cientos de estudiantes, así como

cursos cortos a externos y una gran cantidad de charlas con temas de interés y abiertas a la comunidad.

Las **líneas** principales de investigación han sido varias a lo largo de estos años, concentrándose principalmente en: agricultura de precisión, medicina de precisión, conservación de la biodiversidad y bioinformática; más recientemente se han agregado los temas de astronomía y astrofísica.

Por otro lado, se ha buscado constantemente la reflexión sobre el impacto de la inteligencia artificial (IA) y el aprendizaje automático o *machine learning* (ML) en la sociedad, parte de los temas tratados en los simposios organizados, así como la participación de miembros de PaRMA en múltiples eventos a los que han sido invitados.

El grupo PaRMA ha tenido como objetivo posicionarse a nivel nacional como un actor referente en temas de inteligencia artificial, aprendizaje automático y ciencias de los datos. Además de reportajes en medios nacionales tales como La Nación, el Semanario Universidad, Amelia Rueda y 7 días, PaRMA ha aparecido como referente a nivel país en diversos comunicados del Banco Interamericano de Desarrollo (BID) desde el 2019 y más recientemente con el lanzamiento en el país de la iniciativa fAIr LAC para promover el uso ético de la inteligencia artificial, auspiciado además por el Gobierno de la República.

Comunidades de interés

La mayor parte del trabajo realizado por el grupo ha sido *ad honorem*, tanto por parte de los profesores como de los estudiantes asistentes; algunas investigaciones han sido apoyadas por proyectos inscritos en la Vicerrectoría de Investigación y Extensión (VIE).

Esta forma de trabajo se identificó inicialmente como un modelo de “comunidad de interés”, concepto surgido desde la Escuela de Ingeniería en Computación y apoyado en su momento por el entonces director Mauricio Arroyo Herrera. Este apoyo ha continuado hasta la fecha con el director Roberto Cortés Morales y los compañeros y compañeras que han conformado el Comité Técnico de la Escuela y del Consejo de Escuela de Computación, así como de las vicerrectorías de Investigación y Extensión y de Docencia.

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 unidades académicas 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 multidisciplinarios. Para lograrlo utiliza como motivador la creatividad que surge desde lo interno de la comunidad a partir de los intereses comunes de sus integrantes. (Investiga.TEC; Núm. 35 (2019); 11-13, 1659-3383).

Incluso, en el IV Congreso Institucional del TEC, los profesores Mauricio Arroyo, Saúl Calderón, Laura Coto y Esteban Arias presentamos la propuesta “*Planteamiento de una plataforma de potenciación de actividades extra y cocurriculares por medio de comunidades de interés de innovación*”, la cual fue apoyada por una amplia mayoría de congresistas y aprobada en este Congreso. La propuesta busca dar un reconocimiento y mayor visibilidad a grupos tales como PaRMA, Singular, IEEE

y Habilidades Blandas, en la Escuela de Computación, y otros más en la institución. Esto amplía el manual de cargas para incluir horas dedicadas al trabajo en estos grupos como parte de la jornada docente.

Proyectos recientes

Luego de un primer reporte que se publicó en este mismo medio en mayo de 2019 sobre las actividades del grupo PaRMA, en los últimos años se ha continuado con el desarrollo de varias iniciativas de investigación, algunas formales mediante la VIE y otras *ad honorem*.

En docencia se ha continuado con el apoyo al desarrollo del programa de un año de ciencias de los datos, de la Escuela de Computación, que inició en 2019 y se ha consolidado como un referente a nivel nacional. Coordinado por la

profesora Alicia Salazar, este programa ha tenido una gran demanda. A la fecha se han abierto 16 grupos con un promedio de 25 estudiantes por grupo de profesionales externos. (Más información en: <https://www.tec.ac.cr/fundatec/programa-ciencias-datos-escuela-computacion>).

A nivel de posgrado se ha venido consolidando la impartición de cursos en las áreas de “*machine learning*” y “*deep learning*” también desde el año 2019, donde colaboran profesores miembros de PaRMA. Igual ocurre a nivel de bachillerato mediante las escuelas de verano, cursos cortos y como parte de los contenidos del curso de inteligencia artificial de la carrera de Ingeniería en Computación del TEC.

Generación de un algoritmo de recomendación de métodos para pronósticos a partir de la

caracterización de la serie temporal

Los profesores Martín Solís y Luis-Alexánder Calvo-Valverde trabajan actualmente en este proyecto, planeado para el periodo 2021-2024, en conjunto con investigadores de la Escuela de Administración de Empresas y el investigador externo Hernán Rojas Blanco.

Una serie de tiempo son los valores de una variable en una sucesión cronológica que puede ser anual, trimestral, mensual, diaria, etc. La predicción futura del comportamiento de una serie de tiempo es una tarea que se lleva a cabo en los diferentes campos del conocimiento. Por ejemplo, en epidemiología se realizan predicciones futuras de la incidencia de enfermedades como dengue e influenza; en agronomía se realizan predicciones de la producción futura de un tipo de cultivo; y en meteorología se intenta

Sample of Aniba venezuelana Mez.



Image by Nelson Zamora Villalobos, retrieved from PlantNet.

Groundtruth: Recognized by its oblong leaves, with conspicuously reticulate fine tertiary veins when dry, and a pale brown to reddish brown tone.

Glossary

Oblong: Longer than its width, more or less rectangular.

Reticulate: Shaped like a net.

Ilustración 2. Generación automática de descripciones de fotos de plantas. Proyecto de Adán Mora-Fallas.

predecir el comportamiento futuro de variables climatológicas como radiación solar, humedad relativa, temperatura; y así se pueden seguir dando ejemplos de otros ámbitos.

Al ser una tarea necesaria y común, en diversos campos han surgido muchos métodos y algoritmos para realizar predicciones futuras en series de tiempo, que van desde las clásicas técnicas estadísticas, como los modelos de *Box and Jenkins*, hasta los modelos de aprendizaje automático e inteligencia artificial. De esta manera, cuando en una organización o en una investigación se desea entrenar un modelo para generar pronósticos, hay que lidiar con la tarea de seleccionar los métodos con los que se experimentará entre una amplia gama existente.

Esta investigación pretende generar un nuevo aporte en esa tarea de selección, desarrollando un algoritmo de recomendación mediante meta-aprendizaje, que permita sugerir cuál método es el más conveniente de acuerdo con el problema que se tiene que enfrentar, es decir, de acuerdo con las características de la serie de tiempo y el tipo de pronóstico que desea realizar. Los hallazgos de esta investigación pueden ser utilizados en términos prácticos en ámbitos donde se requieran realizar pronósticos a futuro de una serie de tiempo.

El algoritmo de recomendación que se espera desarrollar puede ponerse al servicio de organizaciones públicas o privadas que requieren hacer pronósticos. Ahora bien, en términos específicos el algoritmo se probará en la elaboración de pronósticos de variables macroeconómicas que se utilizan para la definición de política económica, como son: la inflación, el tipo de cambio y el índice mensual de actividad económica. En términos académicos se espera que el proyecto pueda brindar nuevo conocimiento sobre la capacidad predictiva de los algoritmos según las características de las series de tiempo.

Proyectos de tesis de maestría en Ciencias de la Computación, trabajos finales de graduación y práctica profesional

Actualmente el académico José Carranza, después de haber estado dos años trabajando en Japón, dirige varios proyectos de tesis del Programa de Maestría en Computación del TEC:

- “*Prestando atención a la taxonomía: Uso de jerarquía de clases y atención en la identificación de plantas en un dominio cruzado*”, de Kelvin Jiménez Morales.
- “*Identificación semi-supervisada de plantas: Mejorando la identificación con fotos no etiquetadas*”, de Nicole Carvajal Barboza.
- “*Máquina del tiempo de herbario: Revirtiendo la senescencia de imágenes de herbario usando datos sin pares y pérdidas perceptuales*”, de Esteban Adán Esquivel Barboza.

- “*Alucinador de plantas plausible: Reconstrucción múltiple de imágenes sintéticas de plantas mediante la inyección de ruido*”, de Kevin Quesada Montero.

- “*Generación automática de descripciones de plantas a partir de imágenes*”, de Adán Mora-Fallas.

Los profesores Saúl Calderón y Juan Esquivel Rodríguez dirigieron el trabajo de tesis de Maestría en Computación de Willard Zamora Cárdenas titulada “*Enforcing Morphological Information in Fully Convolutional Networks to Improve Cell Instance Segmentation in Fluorescence Microscopy Images*”.

También los profesores Saúl Calderón y Esteban Arias colaboraron con el proyecto de graduación en Ingeniería en Computación de Danny Xie Li: “*Unlabeled Data for Improving Model Robustness, uncertainty estimation and explainability*”. Además, el profesor Calderón dirigió el proyecto de licenciatura de Ingeniería en Computadoras de Erick Muñoz: “*Efecto de la diferencia entre las distribuciones de datos etiquetados y no etiquetados en la robustez de modelos semi-supervisados*”.

En este 2021 el profesor Arias colaboró con Diego Barquero Morera, de Ingeniería en Biotecnología, como asesor del trabajo final de graduación titulado “*Propuesta de comparación de flujos metabólicos mediante la adaptación de algoritmos para rutas metabólicas: caso práctico con *Chlorella vulgaris* e *Isochrysis galbana**”.



Ilustración 3. Detección y análisis del comportamiento social humano mediante cámaras y sensores.

Detección y análisis del comportamiento social humano mediante cámaras y sensores

Como parte del Laboratorio de Procesamiento de Señales e Imágenes (SIPLab) de la Escuela de Ingeniería Electrónica del TEC, la profesora e investigadora Laura Cabrera-Quirós investiga el uso de múltiples sensores y cámaras en escenarios reales para la detección y análisis del comportamiento social humano para diferentes áreas, desde salud y bienestar humano hasta

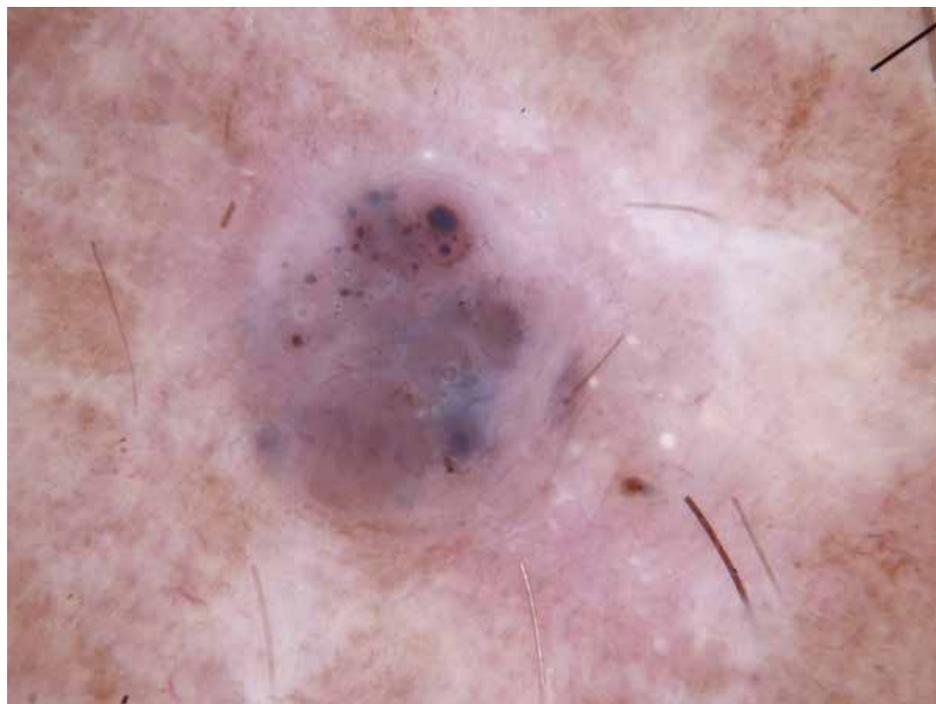


Ilustración 4. Diagnóstico del cáncer de piel tipo melanoma en Costa Rica, mediante el análisis de imágenes y cuantificación del color.

comportamiento organizacional. Y actualmente se encuentra trabajando en el proyecto “Detección de atracción humana en eventos de citas rápidas”.

Este proyecto, en el que trabajan dos estudiantes de la maestría en Electrónica del TEC y que se realiza en colaboración con la Universidad Tecnológica de Delft (Países Bajos), busca detectar y predecir las potenciales respuestas de los participantes sobre la atracción sentida ante cada cita, por medios no intrusivos como sensores de movimiento (*wearables*) y video sin audio. Esta alternativa busca preservar la privacidad de los participantes y su comportamiento natural, mientras abre una ventana a nuevas formas de conocer pareja con apoyo de la tecnología.

Otros proyectos

Por su parte el profesor Luis-Alexánder Calvo-Valverde dirige y ha participado en varios proyectos académicos:

- *Análisis de imágenes y cuantificación del color para el diagnóstico del cáncer de piel tipo melanoma en Costa Rica.* 2020-2021. En conjunto con los investigadores de la Escuela de Física Dionisio Gutiérrez Fallas, Ernesto Montero Zeledón, Mac Arturo Murillo Fernández y José Esteban Pérez Hidalgo.
- *Aprendizaje máquina aplicado al pronóstico en cultivos agrícolas.* DOCINADE. 2013-2019. Tema de tesis doctoral.
- *Investigación sobre un servicio de extensión agrícola utilizando IA.* 2021. Es un proyecto que se desarrolla por medio de la FundaTEC y que se trabaja con la Alianza Bioversity-CiAT en colaboración con CORBANA y el MAG. Por el TEC se cuenta con el apoyo de estudiantes de Ingeniería en Computación. Más información en <https://bigdata.cgiar.org/inspire/inspire-challenge-2020/holatalia-boosting-extension-service-through-ai/>
- *Creación de un instrumento informático para la recopilación de datos radiómicos relacionados con la mamografía como prueba de tamizaje en Costa Rica (MammoTrck).* UCR-ITCR, Costa Rica. 2020. Por el TEC participan los profesores Luis Alexander Calvo, Saúl Calderón Ramírez y la máster Ivonne Madrigal, de la Escuela de Diseño Industrial. Por la Universidad de Costa Rica los académicos Mario Umaña y Sergio Solís.

Prácticas en el extranjero

En los últimos años, varios estudiantes de Ingeniería en Computación y de Ingeniería en Mecatrónica han viajado a diversos países como España, Francia, Canadá y Argentina, gracias a contactos de miembros del grupo PaRMA con

centros de investigación externos. Acá un breve reporte de los proyectos trabajados por los ahora egresados y sus asesores.

En el Grupo Integrado de Ingeniería (GII) de la Universidade da Coruña, España:

- Kenneth Gabriel Laurent Alvarado “*Diseño de un estabilizador inercial para una cámara hiperspectral en un UAV de ala fija*”, asesor profesor Juan Luis Crespo Mariño.
- Eva Corella Solís “*Diseño y desarrollo de un simulador de código abierto para un robot submarino de propósito general*”, asesor Juan Luis Crespo Mariño.
- Ariel Rodríguez Jiménez “*Aplicación de deep learning al aprendizaje de modelos de mundo en mecanismos cognitivos*”, asesor profesor Esteban Arias-Méndez.
- José Antonio Ruiz Jara “*Modificación de librería de control para robot humanoide Poppy*”, asesor profesor Esteban Arias-Méndez.
- Luis Castillo Valverde “*Desarrollo de un experimento de robótica colectiva mediante evolución distribuida*”, asesor profesor Esteban Arias-Méndez.

- Steven Pacheco Portuguez “*Aplicación de desarrollo morfológico al aprendizaje de tareas en dispositivos robóticos*”, asesor profesor Esteban Arias-Méndez.

- Esteban Adán Esquivel Barboza “*Adaptación de algoritmos de visión por computador para su ejecución en Smartphone*”, asesor profesor Esteban Arias-Méndez.

Universitat Rovira i Virgili, Tarragona, España:

- Iván Calvo “*Assessing the impact of a preprocessing stage on deep learning architectures for breast tumor multi-class classification with histopathological images*”, asesor profesor Saúl Calderón.

Universidad Nacional del Litoral, Argentina:

- Erick Alfaro y Ximena Bolanos “*A Brief Analysis of U-Net and Mask R-CNN for Skin Lesion Segmentation*”, asesor profesor Saúl Calderón.

CIRAD, Montpellier, Francia :

- Juan Villacis Llobet “*Cross-domain plant identification*”, asesor profesor Erick Mata.
- Adán Mora-Fallas “*Plants and organs instance segmentation with Mask RCNN*”

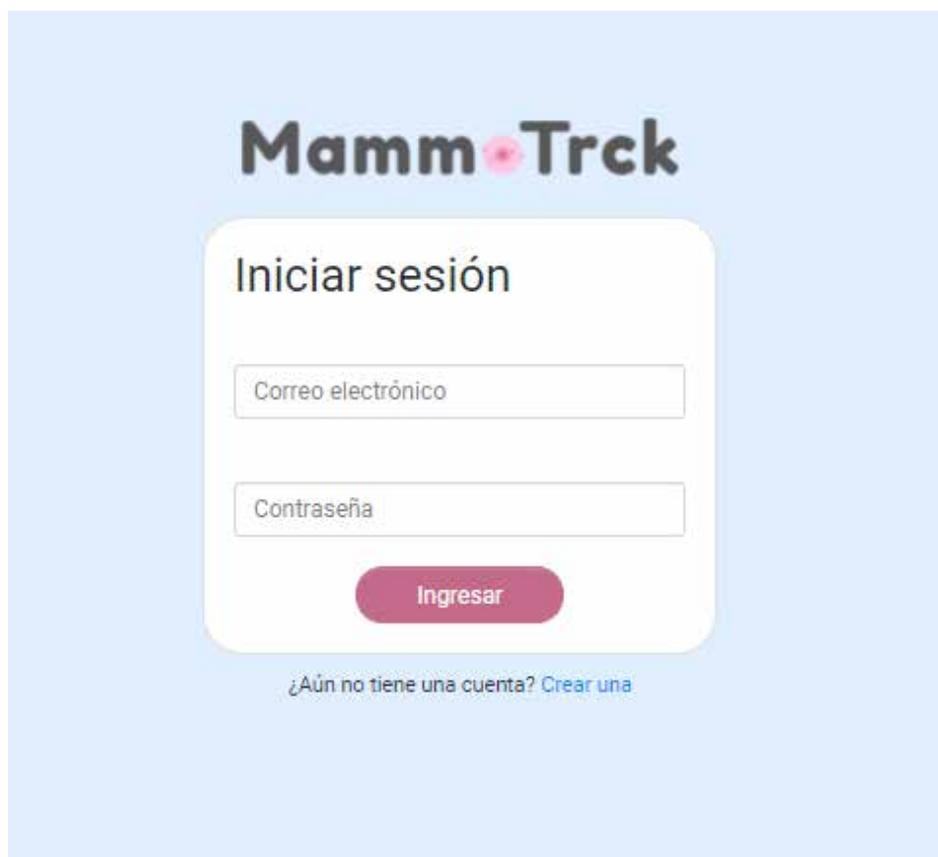


Ilustración 5. Desarrollo del app MammoTrck.

and active learning for massive annotation”, asesor profesor Erick Mata.

Universidad de Toronto, Canadá:

- Mauro Mendez “Using cluster analysis to assess the impact of dataset heterogeneity on deep convolutional network accuracy: A first glance”, asesor profesor Saúl Calderón.
- Ariana Bermúdez “A first glance to the quality assessment of dental photostimulable phosphor plates with deep learning”, asesor profesor Saúl Calderón.

Apoyo a trabajos doctorales y colaboración con eScience

El profesor Erick Mata, quien ha apoyado al grupo PaRMA desde el inicio como profesor miembro del grupo y como coordinador del programa eScience, ha colaborado en varios proyectos.

Junto a Geovanni Figueroa, como parte de su trabajo de tesis doctoral “Automatic Costa Rican Tree Species Identification, Using Deep Learning and Wood Cuts Images”, trabajaron en dos proyectos asociados:

- Identificación automática de especies forestales maderables amenazadas de Costa Rica, mediante técnicas de visión artificial (2017-2019). Y,
- Descubriendo los rasgos usados por algoritmos de Deep Learning en la identificación de taxones de plantas (2019-2020).

Juan Villacis, durante su estancia en Francia participó y ganó la competencia mundial PlantCLEF 2020, desarrollando una herramienta para identificar especies de plantas a partir de imágenes; esto ha sido un tema prominente en las investigaciones de Informática para la Conservación de la Biodiversidad del Programa de Investigación eScience. Además, dio como fruto el artículo “Domain adaptation in the context of herbarium collections. A submission to PlantCLEF 2020”.

Colaboraciones y nuevos temas

De la mano del profesor Saúl Calderón y los

egresados Danny Xie y Erick Muñoz, PaRMA también ha colaborado con el proyecto “ML4H Auditing: From Paper to Practice”, un marco de trabajo propuesto por International Telecommunication Union/World Health Organization ITU/WHO Focus Group on Artificial Intelligence for Health (FG-AI4H).

Otra novedosa área de investigación, que es parte del Laboratorio de Inteligencia Artificial para las Ciencias Naturales (LIANA), es la aplicación de métodos de aprendizaje automático para la resolución de problemas en el área de la astronomía y la astrofísica. En esta área, el profesor e investigador Felipe Meza-Obando con el apoyo de varios estudiantes colaboradores, trabaja en una serie de proyectos en los que se incluye el uso de algoritmos avanzados de aprendizaje automático para la predicción de parámetros estelares complejos de determinar; esto a partir de parámetros más fáciles de obtener mediante sondas espaciales o desde observatorios en Tierra. Este proyecto se desarrolla en colaboración con investigadores de la Universidad de Valencia, España.

Otro proyecto en esta temática consiste en la estimación de la distancia de objetos lejanos, como ciertas galaxias, a partir de datos fotométricos asociados a los índices de color. Adicionalmente y como parte de su investigación doctoral, el profesor Felipe Meza labora en el diseño de un sistema para la predicción del clima espacial, a partir de datos producto de la actividad solar tales como señales de rayos X e imágenes solares; este proyecto es parte de la construcción del primer Radio Observatorio costarricense liderado por el Centro de Investigaciones Espaciales de la Universidad de Costa Rica.

Los próximos cinco años

Actualmente se está trabajando con la organización de la conferencia internacional BIP 2021 que será en Costa Rica y en modalidad virtual, coorganizada por investigadores de todas las universidades públicas del país, donde el grupo ha tenido una participación muy activa. Ver: <https://www.bipconference.org/>

En su V aniversario el grupo PaRMA entrará en una fase de autoanálisis y reflexión de cara al futuro próximo, para continuar colaborando, proponiendo y buscando alianzas con más

colaboradores para el desarrollo de los temas de interés del grupo, para el TEC y para el país.

Se agradece a quienes han sido parte de PaRMA en estos cinco años, quienes han colaborado para el crecimiento y desarrollo de esta comunidad. ■

Colaboraciones

Colaboraron con información para este artículo:

- Dr. Martín Solís, Escuela de Administración de Empresas
- Dr. Luis-Alexánder Calvo-Valverde, Escuela de Ingeniería en Computación
- Dr. José Carranza-Rojas, Escuela de Ingeniería en Computación
- Dra.-Ing. Laura Cabrera-Quirós, Escuela de Ingeniería en Electrónica
- Dr. Erick Mata Montero, Escuela de Ingeniería en Computación
- M.Sc. Saúl Calderón Ramírez, Escuela de Ingeniería en Computación

Contacto

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

***Esteban Arias Méndez** es profesor e investigador de la Escuela de Ingeniería en Computación del TEC. Es máster en Ciencias de la Computación y actual coordinador del grupo de investigación PaRMA para los años 2020 y 2021. Ha colaborado desde sus inicios con el grupo PaRMA en la realización de todas sus actividades. Además, coordina el grupo de voluntariado estudiantil *Singular* y es profesor asesor de los capítulos estudiantiles IEEE Computer Society, CS y Electron Devices Society (EDS); también es profesor consejero de la Rama Estudiantil IEEE del TEC. Participa frecuentemente en actividades de voluntariado y extensión universitaria.

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