



Unidad Monterrey

Cátedras CONACYT 2017

Datos Generales

Institución: Centro de Investigación en Matemáticas, A.C.

Dependencia: CIMAT – Unidad Monterrey

Título / Nombre del proyecto: Desarrollo de algoritmos de cómputo de alto rendimiento para datos masivos complejos.

Número de proyecto: 720

Modalidad: Individual

Tema PECiTI (Ver convocatoria): Desarrollo Tecnológico

Reto PECiTI (Ver convocatoria): Conectividad informática y desarrollo de las tecnologías de la información, la comunicación y las telecomunicaciones

Entidad Federativa donde se realizará el proyecto: Nuevo León

Responsable de la propuesta: Dra. Graciela González Farías

Descripción del Proyecto

Propósito

Dentro de las líneas de investigación que se cultivan en el Centro de Investigación en Matemáticas A.C. (CIMAT) Unidad Monterrey, se encuentran los temas relacionados con el análisis y uso de grandes cantidades de datos de tipo estructurado y no estructurado. Actualmente, se cuenta con investigadores con experiencia en estadística, optimización, métodos numéricos, entre otras. Sin embargo, no se cuenta aún con suficientes especialistas en el desarrollo de métodos computacionales para el uso eficiente de los recursos de infraestructura con los que se cuenta.

Esta propuesta de Cátedras pretende complementar el grupo de investigadores con un profesional que tenga como función el diseño de metodologías y la puesta en marcha de proyectos de alto nivel en el área de Ciencia de Datos, Cómputo Estadístico y/o Cómputo de Alto Rendimiento. Contar con un especialista como el que se solicita contribuirá a lograr las metas a largo plazo en investigación aplicada del CIMAT Monterrey.

Objetivos

Fortalecer el equipo de investigadores para contar con un grupo sólido en cómputo de alto rendimiento, de acuerdo con la visión de largo plazo del CIMAT Unidad Monterrey. La persona seleccionada para cubrir la plaza propuesta deberá convertirse a mediano plazo en líder en temas relacionados al correcto uso de la infraestructura de cómputo de alto desempeño del CIMAT Monterrey y aportará su experiencia en proyectos de investigación y vinculación. Así mismo, participará en



Unidad Monterrey

el programa de maestría en Cómputo Estadístico que se inició en agosto del 2016. La persona que ocupe la Cátedra se incorporará al equipo de investigadores que ya se tiene.

El objetivo a largo alcance es posicionar al CIMAT Monterrey como una institución líder en investigación y docencia, con fuerte influencia en el sector académico y alto reconocimiento en el sector privado a través de proyectos de ciencia aplicada y cómputo de alto rendimiento en el noreste de México.

Motivos de la institución para desarrollar el proyecto

El CIMAT es el único Centro CONACYT en el área de matemáticas y sus aplicaciones, y la Unidad Monterrey es su representación en el noreste del país. Por otro lado, la Ciudad de Monterrey, donde estaría radicada la plaza de Cátedras, tiene una realidad industrial dinámica que demanda de instituciones de investigación innovadoras.

El CIMAT Monterrey planea convertirse en un nodo de investigación aplicada al servicio del sector público y privado que lo rodea, así como en una institución académica que forme recursos humanos de alto nivel para satisfacer las necesidades científico-técnicas de las áreas relacionadas a las matemáticas aplicadas.

Para lograr esto, es necesario mantenerse a la vanguardia en temas de investigación orientados a los retos actuales de cómputo científico y el manejo de grandes cantidades de datos no estructurados, por lo que contar con una plaza de Cátedras en esta área beneficiará mucho a los trabajos del equipo de CIMAT Monterrey.

Contribución esperada de la(s) cátedra(s)

Se espera que con las actividades a desarrollar por el/la nuevo(a) investigador(a) de Cátedras, se vean potenciadas las actividades de investigación en trabajos conjuntos de las diversas áreas de investigación del CIMAT Monterrey como los son el cómputo científico, ciencia de datos y estadística.

Ejemplo de esto serán los trabajos de investigación en cómputo estadístico para grandes volúmenes de datos que implican un amplio conocimiento de las metodologías estadísticas y un manejo selectivo de las estructuras computacionales para analizar eficientemente la información.

A través de las colaboraciones con investigadores del CIMAT e instituciones con las que se colabora se espera generar proyectos conjuntos en los que se

coadyuve a la solución de problemas aplicados en ingeniería, medicina, entre otros, contribuyendo al desarrollo social y productivo del país.

Actividades

Actividades de Investigación:

- Investigación en el área de tecnologías computacionales para el análisis y visualización de grandes volúmenes de datos
- Diseño de metodologías y buenas prácticas para el mantenimiento de bases de datos
- Publicación de reportes técnicos y artículos en revistas indexadas
- Colaboración en proyectos de investigación del equipo de investigadores del CIMAT Unidad Monterrey, las demás sedes del CIMAT, así como con investigadores de otras instituciones académicas nacionales e internacionales
- Exposición en Seminarios, Congresos y Talleres internos y externos
- Ingresar al SNI en los primeros 2 años
- Conseguir fondos para proyectos de investigación

Actividades de Docencia:

- Adscripción al comité de posgrados al que corresponda en el CIMAT
- Apoyo en el fortalecimiento del programa de maestría en Cómputo Estadístico del CIMAT Unidad Monterrey
- Difusión de los programas de posgrado y selección de estudiantes candidatos
- Formación de recursos humanos en programas de posgrado de instituciones de educación superior asociadas y en los demás programas de CIMAT.
- Impartición de cursos como docente
- Dirección y codirección de tesis a mediano plazo en el CIMAT-Monterrey; en el periodo intermedio en los programas en CIMAT-Guanajuato e instituciones académicas con las que el CIMAT tenga convenios

Actividades de Vinculación:

- Participación y organización de Talleres y Congresos tanto de investigación como de vinculación y difusión
- Realización de estancias académicas
- Participación en proyectos de vinculación
- Participación en la red de vinculación del Parque de Investigación e Innovación Tecnológica (PIIT)

- Diseño de soluciones de Software y colaboración con los técnicos académicos para su seguimiento, documentación y entrega en forma de resultados de vinculación

Otras actividades:

- Supervisión del mantenimiento y futuras actualizaciones de la infraestructura de cómputo de alto desempeño del centro
- Establecimiento de protocolos de comunicación entre la infraestructura de cómputo local y la del CIMAT-Guanajuato

Resultados e impactos esperados a 5 y 10 años

En el CIMAT Unidad Monterrey se realiza investigación sobre modelos matemáticos y estadísticos para la generación de información relevante a partir de grandes volúmenes de datos, y la incorporación de un especialista como el que se espera contratar con esta Cátedra busca fortalecer y consolidar a la ciencia de datos como la línea central de investigación.

Los resultados esperados en un plazo de 5 años estarán basados en la contribución y generación de proyectos vanguardistas de investigación en Ciencia de Datos con la producción que esto conlleva: artículos publicados en revistas indexadas, colaboraciones académicas con otros grupos de investigación, presentaciones en conferencias especializadas de nivel internacional.

De manera directa, el nuevo investigador fortalecerá el programa de Posgrado en Cómputo Estadístico con enfoque al análisis de datos cuya primera generación ingresó en agosto del 2016. El investigador contribuirá en la impartición de cursos, así como a la dirección y codirección de tesis.

Se espera el fomento de trabajos de colaboración, reflejados en el inicio de nuevos proyectos de aplicación científico-técnica en las que los conocimientos generados en nuestra institución puedan aprovecharse para resolver problemas aplicados en el sector público y privado de la región noreste del país. Lo anterior mediante el desarrollo de herramientas que puedan ser útiles a la sociedad en general, en particular en áreas de la salud, ciencias económicas y manufactura. En los primeros 5 años de trabajo, se espera que la Cátedra contribuya a establecer colaboraciones y ser una referencia en el uso de las técnicas desarrolladas por nuestro equipo de investigación. Un buen punto de partida son las relaciones que ya se tienen con diversas empresas e institutos de investigación de la región, como la Red de Vinculación del PIIT, e instituciones de investigación y educación superior de la localidad como la UANL, el ITESM, entre otros.



Unidad Monterrey

Una vez alcanzadas las metas de los primeros años, se pretende que la investigación aplicada del CIMAT Monterrey se convierta en un referente como motor de impulso a las actividades económicas de la región noreste del país y ampliar el radio de influencia a otras regiones.

La visión a 10 años es consolidar un grupo de investigación en ciencia de datos de talla internacional, sostener un fuerte ritmo de publicaciones, generación de capital humano especializado y proyectos aplicados con impacto directo en la realidad de nuestro país atendiendo las demandas más prioritarias de los Planes Estatales de Desarrollo, así como los de Ciencia y Tecnología. Por otro lado, se espera contar con un programa de especialización a nivel doctoral, el cual el nuevo investigador apoyará en su creación.

En general, se planea que en un plazo de 10 años se logre un lugar de renombre internacional como grupo de investigación, así como ser una institución reconocida por la constante ejecución de proyectos de ciencia aplicada y vinculación.

Vinculación, pertinencia y perspectivas de transferencia de tecnología (si aplica)

En el CIMAT Unidad Monterrey se tienen diversas interacciones con empresas e instituciones del sector público que generan información con estructuras complejas, entre las que destacan la creación y aplicación de modelos estadísticos multivariados, algoritmos computacionales y herramientas de minería de datos. En aras de fortalecer estas áreas de investigación como pilares que apuntalen al CIMAT como institución líder en modelación estadística de problemas complejos, el CIMAT Unidad Monterrey solicita la integración de una Cátedra CONACYT que complemente al equipo de investigadores que actualmente tiene. El/la nuevo(a) investigador(a) estará temáticamente enfocado(a) en el desarrollo de herramientas de cómputo para el análisis eficiente de grandes cantidades de datos para ser usadas en la aplicación de modelos matemáticos y computacionales a diversas áreas del conocimiento.

Además, el CIMAT Unidad Monterrey ha iniciado recientemente un programa de Maestría en Cómputo Estadístico con enfoque al análisis de datos, uno de los primeros de su tipo en el país, por lo que este nuevo investigador permitirá fortalecer el cuerpo académico del mismo.

Estos dos ingredientes, la generación de capital humano joven en áreas tecnológicas especializadas y la permanente labor de investigación por parte de todo un equipo son lo que dará las condiciones propicias para la transferencia de tecnología hacia los sectores que así lo demanden. Una buena base para esto ya

se tiene, pues el CIMAT Unidad Monterrey fue creado desde sus inicios con una vocación de transferencia de los conocimientos matemáticos con los que el CIMAT cuenta, misma vocación que perdura y que se planea siga siendo el motor de las actividades de la institución.

Descripción del grupo de Investigación o personal académico asociado al proyecto (Incluir los nombres de los miembros y sus especialidades):

El CIMAT Unidad Monterrey actualmente cuenta con un grupo de 7 investigadores enfocados a realizar investigación en temas relacionados a la ciencia de datos y cómputo científico aplicado, el cual se planea fortalecer con la incorporación de la nueva cátedra.

Este equipo es un grupo multidisciplinario con especialistas en estadística, matemáticas y cómputo científico. El grupo de trabajo está liderado por la Dra. Graciela María González Farías (SNI 3, CVU 15311) con especialidad en Series de tiempo y modelos multivariados. El resto de grupo está conformado por el Dr. Jonathan Montalvo Urquizo (SNI 1, CVU 200432) con especialidad en Cómputo Científico y Análisis de Datos, la Dra. María Guadalupe Villarreal Marroquín (SNI 1, CVU 176382) especialista en temas de Optimización y Análisis de Datos, el Dr. Rodrigo Macías Páez (SNI 1, CVU 163691) experto en modelos de escalamiento multidimensional, clustering y clasificación, el Dr. Baidya Nath Saha (SNI 1, CVU 596907) experto en aprendizaje automático e inteligencia artificial, el Dr. Víctor Muñoz Sánchez (CVU 41539) especialista en Análisis de datos de alta dimensión y modelos de clasificación y el Dr. José Jaime Hernández Castillo (CVU 41178) experto en geometría algebraica compleja y análisis de datos, y el Dr. Ritwik Bhattacharya a nivel posdoctoral desarrollando modelos de análisis de supervivencia multivariados.

Además, el CIMAT Monterrey cuenta con personal de apoyo, los llamados técnicos académicos, formado por 7 personas con grados de Maestría y cuya labor es complementar las actividades de investigación y ejecución de proyectos de vinculación.

Descripción de la infraestructura física disponible para ejecutar las actividades del proyecto:

La infraestructura física del CIMAT-Unidad Monterrey disponibles para ejecutar las actividades del proyecto se describen a continuación: El proyecto arquitectónico completo de la Unidad Monterrey del CIMAT en el Parque de Investigación e Innovación Tecnológica consta de tres módulos con 6,255 metros cuadrados de construcción total, y más de 2,671 metros cuadrados de áreas externas. El proyecto incluye biblioteca, cocina, comedor, oficinas administrativas, almacén, taller para equipos de cómputo, sala de videoconferencias, dos salones para

seminarios multimedia, una sala de usos múltiples, una sala de juntas, y oficinas para 32 investigadores, 76 estudiantes, 32 técnicos académicos, 10 directivos (dirección, coordinaciones, gerencias) y 16 auxiliares. En el momento de escribir esta propuesta (febrero 2017) se encuentra terminado y equipado un 44% de la construcción, y un 31% adicional se encuentra en fase de acabado y equipamiento. La obra exterior está prácticamente terminada.

Actualmente el CIMAT Unidad Monterrey cuenta, en materia de infraestructura de cómputo de alto desempeño, con 6 servidores HP DL380 Gen9 8SFF CTO con las siguientes características: 2 procesadores Xeon E5-2640v3 (2.4GHz/8-core), 128GB de memoria RAM (8 x 16GB), 2 discos duros de 600GB de 2.5" Hot Plug con tasa de transferencia de 6Gb/s SAS. En materia de almacenamiento se cuenta con 2 controladoras para tener alta disponibilidad en la comunicación con los servidores (SAN). Cada controladora cuenta con memoria de 8GB de memoria cache FBWC, cada controladora cuenta con 24 y 16 discos duros configurados como un arreglo de discos que fue de 1 Raid 6 con 2 discos virtuales un disco con 12 discos físicos y el otro con 11 discos físicos y 1 disco físico de 60t spare para un total de 30 TB de almacenamiento. Adicionalmente se cuenta con 2 Super-Micro con las siguientes características: 4 Procesador 24 CORES GPU compatible INTEL XEON 5600, 96GB de memoria RAM (6 x 16GB), 4 tarjetas de video NVIDIA C2075 TESLA y 2 discos duros de 2TB para su almacenamiento. En materia de telecomunicación se cuenta con 2 enlaces. El primero es un dedicado de 10 Mbs y otro a 60 Mbs los cuales son a través de fibra óptica con internet II.

El CIMAT Unidad Monterrey forma parte como un nodo más del centro de datos de CIMAT, por lo que se tiene acceso a los recursos informáticos y cómputo con los que cuenta la sede y subsedes del centro (clúster "El Insurgentes", biblioteca virtual, medios multimedia, etc.).

Relación del proyecto con algún laboratorio nacional CONACyT (si aplica)

En el año 2015 se conformó el Laboratorio Nacional de GeoInteligencia, mediante la asociación entre CIMAT y el Centro GEO, cuyo objetivo es consolidarse como un eje de desarrollo tecnológico y de apoyo a la investigación científica en temas relacionados a la geomática, análisis espacial y GeoInteligencia en general. Algunas de las temáticas que se abordan en el laboratorio son análisis geoespacial y territorial de grandes volúmenes de información, modelos de logística y movilidad en tiempo real, entre otras.

Los grupos de investigación que conforman el laboratorio cubren temáticas transversales como estrategias de recopilación y manejo de grandes cantidades de información, análisis de los datos con base en métodos estadísticos y de cómputo, así como la visualización de esta información. La Cátedra solicitada



Unidad Monterrey

busca incluir a alguien que apoye las líneas de investigación que ha dado sustento a la conformación del Laboratorio Nacional de GeoInteligencia.

Relación del proyecto con los programas de Posgrado registrados o susceptibles de pertenecer al PNPC:

Las líneas de investigación a fortalecerse con la incorporación de esta cátedra afectarán directamente al Programa en Ciencias de la Computación (Nivel Internacional del PNPC) del CIMAT Guanajuato y en Programa de Maestría en Cómputo Estadístico (también aceptada en el PNPC) del CIMAT Monterrey. De manera indirecta, los estudiantes de los demás programas de maestría y doctorado de CIMAT (todos en nivel internacional dentro del PNPC) se ven beneficiados por la posibilidad de recibir cursos o ser co-asesorados por miembros del equipo de CIMAT Monterrey. Los estudiantes de todos estos programas y sus respectivos cuerpos académicos también podrán beneficiarse de la integración y desarrollo de proyectos de investigación en ciencia de datos, cómputo estadístico y cómputo de alto rendimiento que se desarrollan en el CIMAT Monterrey.

Además, las instituciones locales con programas de posgrado con las que ya se colabora también tendrán beneficios directos pues los investigadores del CIMAT Monterrey imparten cursos de posgrado en posgrados registrados en el PNPC tanto en los niveles en desarrollo como consolidados, principalmente dentro de la Universidad Autónoma de Nuevo León y el Posgrado Interinstitucional en Ciencia y Tecnología en el que participan los centros de investigación COMIMSA, CIATEC, CIATEJ, CIATEQ, CIDESI, CIDETEQ y CIO.

Indique si se crea, amplía o consolida una nueva línea de investigación:

Esta Cátedra tiene como fin ampliar una línea de investigación a través de la generación estricta de metodologías de computación. La Cátedra tiene como fin contar con un especialista en el manejo de estructuras computacionales de alto rendimiento como cómputo en paralelo, algoritmos distribuidos, sistemas de comunicación a través de diversas plataformas, etc. En este sentido, se generará una nueva perspectiva de una línea de investigación que ya se cultiva y se enriquecerá el posgrado.

¿Cuenta con laboratorio nacional?

SI

Nombre del laboratorio nacional

Laboratorio Nacional enfocado en Imagenología por Resonancia Magnética INB-UNAM en alianza con CIMAT

Laboratorio Nacional de GeoInteligencia en alianza CentroGeo con CIMAT



Unidad Monterrey

¿Tiene PNPC?

SI

Nombre del PNPC

Maestrías en Cómputo Estadístico, Computación y Matemáticas Industriales,
Matemáticas, Probabilidad y Estadística

Doctorados en Ciencias: Computación, Matemáticas, Probabilidad y Estadística

