



# PLAN DE ESTUDIOS

Diplomado Inteligencia Artificial Aplicada para Informática de la Salud

2023 - 2024

cic.ipicyt.edu.mx





















## Diplomado en Inteligencia Artificial para Informática de la Salud

# Rubén López-Revilla, Salvador Ruiz-Correa y José Luis Morán-López

División de Biología Molecular, Grupo de Ciencia e Ingeniería Computacional y Centro de Salud Móvil del Instituto Potosino de Investigación Científica y Tecnológica

San Luis Potosí, S.L.P., Mayo de 2023

#### 1. Justificación

De acuerdo con el *Panorama epidemiológico de las enfermedades no transmisibles en México*, la mayoría de los países de América Latina padecen la más alta prevalencia de enfermedades no transmisibles (ENT) del mundo. Éstas causan cerca de dos millones de defunciones anuales en la región (Alcocer et al. 2021).

En México la tasa de mortalidad por ENT es de 466 por 100,000 habitantes, debido principalmente a que el 71.5% de la población tiene sobrepeso u obesidad, una tasa que es más del doble del promedio mundial (Luciani et al. 2022). El consumo de alcohol per cápita es de 4.4 litros por persona en mayores de 15 años. Además, es muy alta la prevalencia de niveles excesivos de colesterol, consumo de tabaco e hipertensión, agravados por la reducida actividad física.

Las proporciones epidémicas de las ENT tienen como origen los cambios demográficos asociados con el envejecimiento que ocurre en forma sindémica con un crecimiento urbano acelerado y desordenado y los cambios culturales asociados con los hábitos alimenticios en los que predomina el consumo de alimentos altamente procesados, saturados en grasas y azúcares. Los determinantes sociales de la salud como la etnicidad, el género, el nivel educativo y el estado socioeconómico son también elementos conductores de la epidemia de ENT que determinan el acceso a los servicios de salud y sus resultados (Alcocer et al. 2021).

En México las enfermedades transmisibles (ET) también representan una inmensa carga al sistema público de atención médica. Afectan desproporcionadamente a las comunidades con recursos limitados y se vinculan con una compleja variedad de determinantes de la salud superpuestos. La tuberculosis, por ejemplo, tiene una incidencia anual de 23 casos por 100,000















rancisco

habitantes y está entre las 10 primeras causas de mortalidad, después de la COVID-19 y por encima del VIH/Sida (Organización Panamericana de la Salud 2022).

Este panorama epidemiológico demanda la mejoría de las políticas públicas de salud y el desarrollo de nuevas políticas eficaces y rentables para aumentar el acceso y mejorar la calidad de la atención médica que incida sobre los factores de riesgo de las ET y ENT, considerando que la mayoría de la fuerza laboral está desempleada o autoempleada en el sector informal. Está claro que urge reforzar las instituciones de salud con los recursos humanos, tecnológicos y técnicos necesarios para mejorar la atención, el tratamiento y la prevención y reducir el sufrimiento y el costo que representan las principales enfermedades que afectan a la población mexicana.

El Grupo de Ciencia e Ingeniería Computacional y el Centro de Salud Móvil del IPICYT han diseñado el Diplomado en Inteligencia Artificial para Informática de la Salud (DIAIS). El objetivo del programas es contribuir sustancialmente a la formación de recursos humanos capacitados en el manejo de grandes bases de datos y el uso de técnicas de inteligencia artificial que fortalezcan la infraestructura informática médica nacional y desarrollen proyectos aplicados que fortalezcan las capacidades institucionales y promuevan políticas públicas efectivas y rentables para mejorar la atención, la prevención y el tratamiento de las principales enfermedades que aquejan a la población mexicana.

#### Referencias

Alcocer J, López Gatell H, Gallardo MV (2021) Panorama epidemiológico de las enfermedades no transmisibles en México, cierre 2021

Luciani S, Agurto I, Caixeta R, Hennis A (2022) Prioritizing noncommunicable diseases in the Americas region in the era of COVID-19. Rev Panam Salud Pública 46:1. https://doi.org/10.26633/RPSP.2022.83 Organización Panamericana de la Salud (2022) Tuberculosis en la Américas. Informe regional 2021. Washington, D.C.

#### 2. Objetivo del programa

El objetivo principal es formar especialistas en los aspectos teóricos y prácticos del análisis y la ciencia de datos, capaces de recolectar, limpiar, analizar y destilar la información contenida en las bases de datos de las instituciones de salud. La información destilada se utilizará para generar y compartir los conocimientos de manera efectiva para poner a prueba las hipótesis planteadas por especialistas en salud y apoyar a los tomadores de decisiones en el diseño de las estrategias y políticas públicas para mejorar la atención en salud. El programa enfatiza la formación e integración de equipos de trabajo multidisciplinarios formados por ingenieros de datos, analistas y científicos de datos, especialistas en inteligencia de datos para la salud y expertos en ciencias de la salud para desarrollar proyectos orientados a resolver problemas de salud prioritarios de la población mexicana. El Diplomado tendrá una duración cuatro trimestres, y se impartirán anualmente a partir del mes de agosto de 2023.













Los estudiantes regulares que cubran satisfactoriamente los créditos de los cuatro trimestres podrán solicitar y obtener el certificado de Diplomado en Informática de la Salud.

# 3. Metas de los estudiantes participantes

Las metas principales de los estudiantes participante son tres:

- Adquirir los conocimientos básicos que sustentan la teoría y la práctica de la IA.
- Tener la capacidad de utilizar las herramientas básicas en la IA para la solución de problemas prácticos en el área de la salud que requieren el procesamiento de datos y el destilado de información.
- Tener los conocimientos básicos para acceder a un clúster de supercómputo y usar su infraestructura para la resolución de un problema sencillo, pero fácilmente escalable, donde se aplique el aprendizaje profundo.

## 4. Perfil de los aspirantes al diplomado

El diplomado está dirigido a miembros del sistema de salud de México encargados del manejo, uso y destilado de información y conocimiento de las bases de datos de salud de diferentes fuentes y modalidades, tales como historias clínicas electrónicas, resultados de pruebas diagnósticas y exploraciones médicas.

#### 5. Plan de estudios

El plan de estudios incluye conceptos fundamentales para el análisis y la ciencia de datos de la salud, así como metodologías de aprendizaje de máquina, aprendizaje profundo y otras técnicas de inteligencia artificial.

## Primer trimestre (48 h)

Panorama de la Inteligencia Artificial para la Informática de la Salud (4 h)

1. Aplicaciones del aprendizaje de máquina y la inteligencia artificial

## Introducción a la Programación: Conceptos básicos de R (32 h)

- 1. Scripts.
- 2. RStudio.
- 3. Lectura/escritura de datos.
- 4. Tipos de variables.
- 5. Objetos y funciones.
- 6. Operador Pipe.
- 7. Operadores para la filtración de datos.
- 8. La función de combinación c().













- 9. Valores faltantes.
- 10. Creación de columnas.
- 11. Funciones, operadores y ciclos.
- 12. Métodos para sintetizar datos.
- 13. Graficado de datos.
- 14. Ajustes de gráficas.

## Fundamentos de Probabilidad y Estadística para la Informática de la Salud (16 h)

- 1. Axiomas de probabilidad.
- 2. Variables aleatorias.
- 3. Distribuciones de probabilidad.
- 4. Momentos de variables aleatorias.
- 5. Estimación de máxima verosimilitud.

## Segundo trimestre (48 h)

## Estadística Médica Aplicada (32)

- 1. Estadísticas descriptivas.
- 2. Distribuciones de muestreo.
- 3. Estimación de parámetros.
- 4. Pruebas de hipótesis.
- 5. Análisis de varianza.
- 6. Análisis de frecuencias.
- 7. Pruebas no paramétricas.

## Narrativa para las Ciencias de Datos (16)

- 1. Investigación y comunicación científica.
- 2. Publicaciones científicas
- 3. El artículo de investigación
  - a. Formato IMRAD.
  - b. Redacción.
  - c. Publicación.
- 4. Taller de análisis de artículos.

#### Tercer trimestre (48 h)

#### Aprendizaje de Máquina para la Informática de la Salud (48 h)

- 1. Conceptos básicos de aprendizaje de máquina.
- 2. Análisis exploratorio de datos.
- 3. Regresión lineal múltiple.
- 4. Métodos de regresión no lineal.
- 5. Métodos de clasificación.
- 6. Métodos de reducción de dimensionalidad.













7. Introducción al aprendizaje profundo.

## Cuarto trimestre (48 h)

#### Inteligencia Artificial para Informática de la Salud

- 1. Conceptos básicos de aprendizaje profundo.
- 2. Redes neuronales convolucionales.
- 3. Redes neuronales recursivas.
- 4. Autocodificadores
- 5. Aplicaciones de aprendizaje profundo para la Informática de la Salud.

## 6. Requisitos de graduación

Para obtener el certificado de Diplomado en Informática de la Salud, los alumnos deberán:

- 1. Aprobar cada una de las asignaturas trimestrales con una calificación igual o mayor a 8.0.
- Diseñar un Proyecto de Investigación que aborde el análisis de una base de datos relacionada con un problema de salud prioritario en México y entregarlo al final del segundo trimestre.
- 3. Entregar el informe técnico del avance del proyecto al término del tercer trimestre.
- 4. Entregar el informe final del proyecto al término del cuarto trimestre.
- 5. Presentar y defender satisfactoriamente el informe final.

## 7. Instructores, asesores y asistentes

El programa de estudios será impartido por investigadores adscritos al Grupo de Ciencia e Ingeniería Computacional y a la División de Biología Molecular del IPICYT, así como a otras instituciones nacionales e internacionales que forman parte de la Alianza Multidisciplinaria de Especialistas en Salud del Centro de Salud Móvil del IPICYT (https://mhc.ipicyt.edu.mx/mobilehealthcenter/).

## Instructores

Rubén López-Revilla. Es profesor de la División de Biología Molecular (DBM) del IPICYT, Médico por la Universidad Autónoma de San Luis Potosí, Doctor en Genética por el Centro de Investigación y de Estudios Avanzados. Es miembro del Sistema Nacional de Investigadores Nivel 2. Fundó la DBM-IPICYT. Sus intereses de investigación son el diagnóstico y la epidemiología molecular de las enfermedades infecciosas y el cáncer y el monitoreo de la calidad del aire interior para prevenir COVID-19 y otras enfermedades de transmisión aérea. En colaboración con Salvador Ruiz Correa inició el Centro de Salud Móvil, grupo multidisciplinario













de investigación-acción que aborda problemas prioritarios de salud pública con herramientas de informática y aplicaciones móviles.

Ana Paulina Ponce Tadeo. Estudió la licenciatura en Física y la Maestría en ciencias en la Universidad Autónoma de San Luis Potosí, Doctora en Nanociencias y Materiales por el IPICYT. Su línea de investigación se centra en las propiedades estructurales de pequeños agregados magnéticos de la serie 3d y 4d usando cálculos de primeros principios y teoría de Grafos. Además, ha sido docente de nivel superior por 6 años en la universidad Politécnica de San Luis Potosí donde imparte clases de Matemáticas. Actualmente es posdoctorante en la División de Materiales Avanzados y es miembro de Grupo de Ciencia e Ingeniería de Computacionales del IPICYT.

Cesare Moisés Ovando Vázquez. Licenciatura en Ingeniería Física y doctorado en Ciencias con especialidad en Física por el Cinvestav Zacatenco. Es miembro del Sistema Nacional de Investigadores Nivel II. Su investigación es en Bioinformática e Inteligencia Artificial (IA) aplicada al estudio de ARNs cross-kingdom y del síndrome metabólico (MetS). Actualmente dirige el laboratorio de Bioinformática e IA (AI-BioLab) CNS-IPICYT.

Salvador Ruiz Correa. Investigador del CNS. Licenciatura y maestría en Ingeniería Mecánica y Eléctrica por la Universidad Nacional Autónoma de México y doctorado en Ingeniería Eléctrica por la Universidad de Washington en Seattle. Es miembro del Sistema Nacional de Investigadores Nivel I desde 2008. Sus intereses de investigación se centran en aplicaciones de computación móvil, inteligencia artificial y estadística aplicada en proyectos de impacto social en México. Dirige el Youth Innovation Laboratory (You-i Lab) del IPICYT.

Daniel Ignacio Salgado Blanco. Licenciatura en Física y doctorado en Ciencias e Ingenierías de Materiales en la Universidad Nacional Autónoma de México. Actualmente se encuentra laborando como catedrático CONACYT adscrito al CNS-IPICYT. Sus intereses de investigación están ligados al estudio de sistemas coloidales mediante simulaciones moleculares, utilizando computación de alto rendimiento. Es miembro del Sistema Nacional de Investigadores Nivel I.

Mishael Sánchez Pérez. Ingeniero en Computación y Redes de Computadoras por la Universidad Morelos de Cuernavaca y Doctorado directo en Ciencias Computacionales por el ITESM. Su investigación se centra en el análisis de datos masivo utilizando Machine Learning y herramientas bioinformáticas.

#### Asesores

Sinhué López Moreno. Es investigador Catedrático CONACYT adscrito a la División de Materiales Avanzados del IPICYT. Es Ingeniero Electromecánico por el Tecnológico de Ocotlán Jalisco. Realizó los estudios de Maestría y Doctorado en Ciencias de Materiales en el Cinvestav Querétaro. Hizo un posdoctorado en la Universidad de La Laguna (Tenerife, Islas Canarias, España) y otro en la Facultad de Ciencias en la UNAM. Es miembro del Sistema Nacional de













Investigadores con Nivel II. Actualmente desarrolla investigación teórica en materiales en sólidos tridimensionales y de baja dimensionalidad utilizando cálculos de primeros principios.

José Luis Morán López. Realizó sus estudios de doctorado en la Universidad Libre de Berlín. Actualmente es Investigador Titular del Departamento de la División de Materiales Avanzados y Coordinador Académico del Centro Nacional de Supercómputo del Instituto Potosino de Investigación Científica y Tecnológica. Cuenta con un número artículos científicos en revistas y en libros de difusión internacional que asciende a 229. En 1984 se hizo acreedor a la beca de investigación "John Simon Guggenheim", otorgada en Nueva York. En 1988, la Organización de Estados Americanos con sede en Washington, le otorgó el premio "Manuel Noriega Morales", en Ciencias Exactas. Se hizo merecedor, en 1990, al Premio Internacional C. V. Raman del International Centre for Theoretical Physics. En 1996 recibió el Premio Nacional de Ciencias y Artes en el área de Ciencias Físico-matemáticas y Naturales, máxima distinción a un científico en nuestro país. Fue electo Vicepresidente de la Academia Mexicana de Ciencias (2014-2017) y ocupó el puesto de Presidente de Julio de 2017 a Junio de 2020. Su más reciente distinción es la de Investigador Emérito del Sistema Nacional de Investigadores. En sus contribuciones al desarrollo de la investigación científica en nuestro Estado, destacan la fundación del Consejo Potosino de Ciencia y Tecnología en 1996, la creación del Instituto Potosino de Investigación Científico y Tecnológico (2000) y el Centro Nacional de Supercómputo (2006).

Juan Carlos Rosas Cabrera es el Coordinador General del CNS. Egresado de la UAM Iztapalapa de Ing. Electrónica en Computación y con Maestría de Gestión de la Información y Estudios Organizacionales de la UAM Xochimilco y la Universidad del la Habana, Cuba. Fue Coordinador Técnico del Laboratorio de Supercómputo de UAM Iztapalapa donde se desempeñó como responsable de una de las infraestructuras más importantes de México; también participó y estuvo como responsable de proyectos interinstitucionales como la creación de la Delta Metropolitana o LANCAD entre la UNAM, UAM y Cinvestav; con otros centros de Supercómputo de Universidades y Centros Públicos de Investigación de CONACyT en México. También colaboró en proyectos de Supercómputo en otros países como Chile, Italia y España. Participó en proyectos importantes o iniciativas de NIBA o México Conectado de la SCT y Gobierno Federal, en comités de Ciencia y Tecnología y Digital Nacional. Creó uno de los más importantes semilleros de recursos humanos en HPC en México. Fue Director de operaciones en Atalait antes Sungard Latam en soluciones de DRP y BCP en el sector financiero en México. Estuvo a cargo de servicios y soporte como Director en la BUAP en la DCyTIC en todas las dependencias de esta Universidad en el estado de Puebla para satisfacer las necesidades académicas y de investigación del quehacer institucional; así como también colaboró en proyectos de consolidación de servicio y atención del LNS de la BUAP. Estuvo también desempeñándose como Gerente Nacional de la Dirección de Educación del área comercial de Totalplay Empresarial de Grupo Salinas otorgando soluciones de conectividad a gran escala para todo el sector educativo y de investigación en el país. Actualmente el Mtro. Juan Carlos Rosas es el Coordinador General del CNS del IPICYT en donde dirige esfuerzos para dar servicios de HPC, para la Academia y proyectos Sustentables; así como también la generación del recurso humano especialista en soluciones TIC de alta especialidad y grandes capacidades a través de













la innovación tecnológica. En el CNS actualmente se llevan a cabo proyectos importantes con el Sector Salud, Educación, Industria, Investigación y Gobierno.

#### **Asistentes**

Viridiana del Carmen Robledo Valero es Licenciada en Ciencias Ambientales por la Universidad Autónoma de San Luis Potosí y Maestra en Salud Pública por el Instituto Nacional de Salud Pública (INSP, México). Actualmente es estudiante de doctorado del INSP y directora ejecutiva del You-i Lab. Se enfoca en administrar proyectos de salud pública que abordan problemas de salud ambiental y desarrollar plataformas digitales de vigilancia epidemiológica automatizada para la evaluación integrada de riesgos a la salud en comunidades mexicanas.

María de Lourdes Mendoza Flores es Licenciada en Ciencias Ambientales por la Universidad Autónoma de San Luis Potosí y Maestra en Salud Pública por el Instituto Nacional de Salud Pública (INSP, México). Su trabajo se centra en la identificación de la distribución y determinantes de salud mediante metodología epidemiológica para la toma de decisiones.

Rubicel Trujillo Acatitla. Licenciado en Biología por la Universidad Autónoma Metropolitana Unidad Xochimilco. Especialista en Geomática por el Centro de Investigación en Ciencias de Información Geoespacial (Centro GEO). Maestro en Ciencias Ambiental por el Instituto Potosino de Investigación Científica y Tecnológica A.C. (IPICYT). Actualmente, estudiante de Doctorado en la División de Geociencias Aplicadas (IPICYT) enfocado al procesamiento de imágenes satelitales para la detección y segmentación de superficies mediante técnicas de aprendizaje de máquina y aprendizaje profundo.

Daniel Rafael Saldaña Torres es licenciado en Biotecnología Genómica por la Universidad Autónoma de Nuevo León, Facultad de Ciencias Biológicas. Monterrey, Nuevo León, es Maestro en Ciencias en Biología Molecular emitido por el Instituto Potosino de Investigación Científica y Tecnológica, A.C. San Luis Potosí, San Luis Potosí. Actualmente se cursa el 4to semestre de Doctorado en Ciencias en Biología Molecular en el Instituto Potosino de Investigación Científica y Tecnológica, A.C. San Luis Potosí, San Luis Potosí.

#### 8. Costos

El costo del diplomado debe estar cubierto en su totalidad para la emisión de los respectivos reconocimientos. El costo y las facilidades de pago por módulo que se ofrecen son las siguientes:

- La colegiatura por los cuatro trimestres del diplomado en un sólo pago es de MXN \$40,000.00
- La colegiatura trimestral del diplomado con pagos es de MXN \$10,000.00
- La colegiatura por mensualidad del diplomado con pagos es de MXN \$3,350.00













Se recomienda realizar el(los) pago(s) con tarjeta de crédito o débito en el portal de pagos online del IPICYT. Alternativamente, los pagos pueden realizarse mediante depósito de efectivo o cheques, o bien mediante transferencia bancaria a la cuenta que el IPICYT habilitará para tales fines. La información de la cuenta bancaria se enviará tras haberse realizado el registro de la inscripción.

#### 9. Duración

Los módulos constan de dos sesiones a la semana con duración de dos horas cada una. El diplomado tendrá una duración total de 192 horas de clases; lo que corresponde a 12 meses.

## 10. Metodología de enseñanza-aprendizaje y evaluación

El diplomado se imparte a distancia utilizando la plataforma de enseñanza Moodle del IPICYT y una plataforma para videoconferencias. La metodología de enseñanza-aprendizaje considera clases remotas enfocadas en los contenidos de los programas. Las clases audiovisuales se imparten en vivo (aprendizaje sincrónico) y son grabadas para que los estudiantes que no puedan atender a la clase tengan la oportunidad de aprender los contenidos posteriormente (aprendizaje asíncrono). Los participantes cuentan con la asesoría del instructor.

Las clases son complementadas con búsquedas bibliográficas por parte de los participantes, trabajos prácticos y exámenes. Cada asesor dará a conocer el procedimiento de evaluación al inicio de su curso, requiriéndose que cada módulo sea aprobado para acceder al siguiente. La escala de calificación es de 0 a 10 puntos, donde la calificación mínima aprobatoria es 7 según siguiendo los criterios del Reglamento de Posgrado del IPICYT.

## 11. Reconocimientos y calificaciones finales

Los estudiantes que concluyan satisfactoriamente el diplomado recibirán un diploma oficial de participación y una constancia. Las constancias son digitales, se enviarán al correo electrónico registrado por cada participante al momento de su inscripción. Las constancias son foliadas, indican la calificación final obtenida por el participante y la cantidad de horas de especialización acreditadas. Tanto la calificación final como la cantidad de horas de capacitación acreditadas son asentadas en un acta oficial que es entregada al Departamento del Posgrado del IPICYT.













### 12. Proceso de admisión

Los interesados en inscribirse deben llenar el formulario disponible en la página de web del GCIC (<a href="https://cic.ipicyt.edu.mx/">https://cic.ipicyt.edu.mx/</a>) en fechas que la convocatoria se encuentre abierta (verificar la convocatoria para esté punto) indicando:

- 1. Nombre completo del participante. Se deben incluir todos los nombres y apellidos del participante.
- 2. Correo(s) electrónico(s) de contacto. Incluir una o más direcciones de correo electrónico válidas para el envío de la información vinculada al posgrado.
- 3. Número de teléfono. Indicar un número telefónico de contacto a diez dígitos (Incluir lada). Se recomienda enviar número de telefonía celular y/o fija.
- 4. Datos de facturación. Todos los ingresos del IPICYT se facturan. Si el participante requiere factura, debe indicar la RAZÓN SOCIAL, RFC y DIRECCIÓN FISCAL (calle, número exterior e interior, colonia, código postal, ciudad y estado). En cualquier otro caso, el IPICYT emitirá una factura a nombre de "Público en general" para cumplir con sus obligaciones fiscales.
- 5. Modalidad de pago. Indicar si el pago se realizará en una sola exhibición o en parcialidades sin intereses.
- 6. Para cualquier duda podrá comunicarse al correo electrónico atencion.eiais@ipicyt.edu.mx.

