

# Diplomado en Inteligencia Artificial para Informática de la Salud

Plan de estudios



## Diplomado en Inteligencia Artificial Aplicada para Informática de la Salud

### Plan de Estudios

#### Módulo 1. Narrativa de Datos (16 h)

##### **1.1 Protocolos de los proyectos de investigación en salud basados en IA**

Se revisará la estructura y contenido de los protocolos para la planeación de proyectos de investigación en salud basados en IA y la elaboración de los informes de resultados correspondientes.

##### **1.2 Proyectos del diplomado**

Se elaborará una relación con los detalles básicos de los proyectos que incluya a los participantes (estudiantes y asesores internos y externos), así como las responsabilidades, tareas y productos esperados de cada uno.

##### **1.3 Artículo de revisión sobre el problema de salud de cada proyecto**

El propósito del artículo de revisión es analizar la situación actual del problema de salud de cada proyecto, para identificar las limitaciones de los estudios e intervenciones previas que permitan definir y justificar los objetivos, métodos y resultados del proyecto.

El artículo se basará en una plantilla con el formato de un preprint que podría ser depositado en el repositorio del DIAIS (servidor de la Biblioteca del IPICYT) para su consulta en línea.

Se guiará a los estudiantes en el uso de un gestor de referencias, así como en la búsqueda bibliográfica y la selección y anotación de las referencias pertinentes para la redacción del artículo.

##### **1.4 Informes terminales de los proyectos**

La estructura y contenido de los informes terminales se basará en una plantilla con formato similar a la del artículo de revisión.



## Módulo 2. Programación en R para IA (16 h)

### 2.1 Introducción a R

R es el lenguaje de programación y RStudio es uno de los entornos de desarrollo integrados (IDE) más populares para trabajar con R. Se pueden proporcionar lecciones sobre cómo instalar R, RStudio y una breve introducción a la sintaxis básica de R.

### 2.2 Fundamentos de R para Ciencia de Datos e Inteligencia Artificial

Esta sección cubre los fundamentos de la programación en R, tipos de datos, estructuras de control, funciones, entre otros conceptos básicos necesarios para trabajar en ciencia de datos e inteligencia artificial.

### 2.3 Manipulación de Datos con tidyverse (dplyr, tidyr) y data.table

La manipulación de datos en R puede realizarse principalmente con el paquete tidyverse, que incluye las librerías dplyr y tidyr, similares a NumPy y Pandas en Python. Además, data.table es otra opción popular para manipulación eficiente de datos en R.

### 2.4 Desarrollo de Modelos con caret

Caret es una librería en R que proporciona una interfaz unificada para entrenar y comparar modelos de aprendizaje de máquina. Al igual que scikit-learn en Python, caret permite construir y ajustar modelos predictivos de manera eficiente.

### 2.5 Construcción de Redes Neuronales con TensorFlow y Keras (en R)

TensorFlow y Keras también están disponibles en R y se pueden utilizar para construir y entrenar redes neuronales de manera similar a como se hace en Python. La librería keras en R proporciona una interfaz para trabajar con Keras, y TensorFlow puede ser utilizado directamente también.





## Módulo 3. Aprendizaje de Máquina (32 h)

### 3.1: Conceptos Básicos de Aprendizaje de Máquina

Esta establece los pilares fundamentales del Aprendizaje de Máquina, proporcionando una visión amplia que abarca desde la definición del campo hasta sus aplicaciones prácticas en diversos sectores. Además, se introduce a los estudiantes en los conceptos esenciales como datos, rasgos, etiquetas y modelos, sentando las bases para comprender las técnicas de evaluación de modelos.

### 3.2 Aprendizaje Supervisado

Esta sección se centra en el aprendizaje supervisado, donde se presentan técnicas para resolver problemas de regresión y clasificación. Se estudian algoritmos como regresión lineal, regresión logística, árboles de decisión, SVM y redes neuronales, junto con métodos de ensamble como Gradient Boosting y Bagging, que permiten mejorar el rendimiento del modelo.

### 3.3: Aprendizaje No Supervisado

En esta sección se exploran los métodos de aprendizaje no supervisado, utilizados cuando no se disponen de etiquetas en los datos de entrenamiento. Se examinan técnicas de aglomeración como K-Medias y DBSCAN, así como métodos de reducción de la dimensionalidad como PCA y t-SNE, que facilitan la comprensión y visualización de conjuntos de datos complejos.

### 3.4: Aprendizaje por Refuerzo

Se introduce el aprendizaje por refuerzo, un enfoque donde un agente aprende a través de la interacción con un entorno. Se exploran conceptos como los procesos de decisión de Markov (MDP) y se presentan algoritmos básicos de aprendizaje por refuerzo.

## Módulo 4. Aprendizaje Profundo (24 h)

### 4.1: Introducción al Aprendizaje Profundo

Aquí se proporciona una introducción a los conceptos básicos de las redes neuronales profundas, así como una visión general de la evolución y las aplicaciones del aprendizaje profundo en diversos campos como la visión por computadora, el procesamiento de lenguaje natural y la bioinformática.



#### **4.2: Regularización en Redes Neuronales**

En esta sección se aborda el problema del sobreajuste en las redes neuronales y se presentan técnicas de regularización, como dropout y regularización L1 y L2, que ayudan a evitar que el modelo se ajuste demasiado a los datos de entrenamiento. También se exploran métodos para prevenir la explosión y el desvanecimiento del gradiente durante el entrenamiento.

#### **4.3: Métodos de Optimización para Redes Neuronales Profundas**

Aquí se estudian los métodos de optimización utilizados para ajustar los parámetros de las redes neuronales profundas de manera eficiente. Se cubren desde el gradiente descendente y sus variantes hasta técnicas avanzadas como SGD con momentum, RMSProp y Adam, junto con estrategias para la inicialización adecuada de los pesos de la red.

#### **4.4: Redes Neuronales Convolucionales (CNN)**

Esta sección se enfoca en las redes neuronales convolucionales (CNN), una arquitectura especialmente diseñada para procesar datos de tipo imagen. Se exploran los principios de funcionamiento de las CNN, así como arquitecturas comunes como LeNet, AlexNet, VGG y ResNet, y se presentan aplicaciones prácticas en clasificación de imágenes, detección de objetos y segmentación semántica.

#### **4.5: Redes Neuronales Recurrentes (RNN) y Análisis de Series Temporales**

Se introduce a las redes neuronales recurrentes (RNN) y se explora su capacidad para modelar dependencias temporales en datos secuenciales. Se analizan aplicaciones en análisis de series temporales, predicción y generación de secuencias, destacando su utilidad en áreas como finanzas, clima y procesamiento de señales.

#### **4.6: Redes Neuronales Recurrentes en Procesamiento de Lenguaje Natural (NLP)**

En esta última se aborda el uso de las redes neuronales recurrentes en el procesamiento de lenguaje natural (NLP). Se examina el modelado del lenguaje con RNN y se presentan aplicaciones en traducción automática, generación de texto, análisis de sentimientos y reconocimiento de voz, demostrando el potencial de estas técnicas en la comprensión y generación de lenguaje humano.

### **Módulo 5. Seminario de proyecto terminal (16 h)**

