

## Diplomado en Ciencia de Datos con Python 2023

### Introducción

El Diplomado en Ciencia de Datos con Python (DCDP) es una formación profesional cuyo objetivo es proporcionar las bases teórico-prácticas de las disciplinas que integran la ciencia de datos, utilizando el lenguaje de programación Python. Está dirigido a profesionales, académicos o estudiantes, de licenciatura o posgrado, con conocimientos básicos en matemáticas y programación estructurada, que deseen aprender a utilizar herramientas avanzadas de programación en Python para el análisis de datos estructurados y no estructurados.

Se busca que el o la participante logre al concluir el diplomado:

- Usar y configurar la plataforma en la nube COLAB de Google.
- Aplicar la programación vectorizada de Python usando los paquetes y módulos de numpy, pandas, matplotlib, y seaborn.
- Elegir métodos de exploración descriptiva de datos o modelos predictivos de regresión o clasificación en función del problema.
- Desarrollar soluciones propias a problemas mediante métodos de análisis estadístico, aprendizaje automático y aprendizaje profundo.
- Aplicar los paquetes de scikitlearn y keras-tensorflow para construir modelos predictivos.
- Evaluar la calidad de un modelo de clustering, regresión o clasificación.

### Calendario y organización

El diplomado se imparte los días jueves de 17:00 a 20:00 y los sábados de 9:00 a 12:00. Los módulos del diplomado se llevarán a cabo en las siguientes fechas:

#### CALENDARIO DCDP 2023

	ENE		FEB				MAR				ABR				MAY				JUN				JUL			
JUEVES																										
17:00 -20:00	26	2	9	16	23	2	9	16	23	30	6	13	20	27	4	11	18	25	1	8	15	22	29	6	13	20
SABADO																										
9:00-12:00	28	4	11	18	25	4	11	18	25	1	8	15	22	29	6	13	20	27	3	10	17	24	1	8	15	22
MÓDULO	MÓDULO 1									MÓDULO 2									MÓDULO 3							
	Python y Estadística de datos									Machine Learning									Deep Learning							
DURACIÓN	54 HORAS									54 HORAS									42 HORAS							

Código de colores

	Sesión teórico-práctica
	Sesión de presentación de trabajos
	Asueto

## Módulos y temas

El diplomado es una formación teórico-práctica que busca construir habilidades y competencias en el análisis de datos. Por ello, cada módulo está dividido en dos partes: *curso* y *proyecto*. En el curso se imparten los temas teóricos que se refuerzan mediante ejercicios en clase y tareas. En el proyecto se busca que el/la participante demuestre que ha adquirido los conocimientos necesarios para resolver un problema específico. En esta segunda parte se evaluará al alumno/a para otorgarle una calificación.

<p><b>I. Python y Estadística de Datos</b> (Curso: 7 semanas; Proyecto: 2 semanas)</p> <ol style="list-style-type: none"> <li><b>Introducción a Python.</b> <ol style="list-style-type: none"> <li>Variables y tipos de datos</li> <li>Control de flujo de programa</li> <li>Funciones, Clases y Objetos</li> </ol> </li> <li><b>Paquetes para ciencia de datos</b> <ol style="list-style-type: none"> <li>Cálculo científico</li> <li>Visualización de datos</li> </ol> </li> <li><b>Estadística descriptiva</b> <ol style="list-style-type: none"> <li>Cálculos estadísticos</li> <li>Simulación numérica</li> </ol> </li> <li><b>Teoría muestral</b> <ol style="list-style-type: none"> <li>Tipos de distribución</li> <li>Significancia estadística</li> <li>Prueba de hipótesis</li> </ol> </li> </ol> <p><b>II. Machine Learning</b> (Curso: 7 semanas; Proyecto: 2 semanas)</p> <ol style="list-style-type: none"> <li><b>Introducción al ML</b> <ol style="list-style-type: none"> <li>¿Qué es ML?</li> <li>Función de costo</li> <li>Gradiente descendente</li> <li>Regresión lineal</li> </ol> </li> </ol>	<ol style="list-style-type: none"> <li><b>Preprocesamiento de datos</b> <ol style="list-style-type: none"> <li>Limpieza</li> <li>Normalización</li> <li>Transformación</li> <li>Imputación</li> </ol> </li> <li><b>Clasificadores básicos</b> <ol style="list-style-type: none"> <li>Métricas de rendimiento</li> <li>Árboles de decisión</li> <li>Regresión logística</li> <li>SVM</li> <li>K-Nearest Neighbors</li> <li>Bayes Ingenuo</li> </ol> </li> <li><b>Agrupamiento (clustering)</b> <ol style="list-style-type: none"> <li>Medidas de distancia</li> <li>K-MEANS</li> <li>Métodos aglomerativos</li> <li>Reducción de dimensionalidad</li> </ol> </li> </ol> <p><b>II. Deep Learning</b> (Curso: 5 semanas; Proyecto: 2 semanas)</p> <ol style="list-style-type: none"> <li><b>Introducción al Deep Learning</b></li> <li><b>Redes MLP</b></li> <li><b>Redes de Convolución</b></li> <li><b>Auto-encoders</b></li> <li><b>LSTM</b></li> </ol>
--	---

## Evaluación de cada módulo

Cada módulo será evaluado por el instructor en turno mediante la realización de un proyecto. Para ser evaluado se solicitará a cada participante:

- Una presentación
- Un programa
- Un reporte

La evaluación se hará en base a los siguientes criterios:

## RÚBRICA DE EVALUACIÓN DEL PROYECTO

CRITERIO	VALORACIÓN			
	10	9	8	7
<b>PRESENTACIÓN</b>				
1. OBJETIVOS	Los objetivos se exponen con claridad. Se identifica el problema a resolver de manera correcta y las metas a lograr se exponen con precisión.	Los objetivos se exponen con cierta claridad. Se dificulta un poco la identificación del problema a resolver o las metas a lograr carecen de precisión.	Los objetivos se exponen con cierta claridad. Se dificulta bastante la identificación del problema a resolver o las metas a lograr carecen de precisión.	Los objetivos no son claros. Se dificulta mucho la identificación del problema a resolver y las metas a lograr carecen de precisión.
2. METODOLOGÍA	La metodología se expone con claridad. Se identifican todos los pasos correctos a seguir y se presentan en el orden correcto.	La metodología se expone con cierta claridad. Se identifican la mayoría de los pasos correctos a seguir y se presentan en el orden correcto.	La metodología se expone con cierta claridad. Se identifican pocos pasos correctos a seguir o se presentan en el orden incorrecto.	La metodología se expone con poca claridad. Se dificulta mucho la identificación de pasos correctos a seguir y el orden correcto.
<b>PROGRAMA</b>				
1. CÓDIGO	La notebook se muestra ordenada y legible. Se aprovechan de manera óptima los recursos que ofrece la plataforma.	La notebook se muestra ordenada y legible. Se aprovechan de manera suficiente los recursos que ofrece la plataforma.	La notebook se muestra poco ordenada o legible. Se aprovechan pocos recursos que ofrece la plataforma.	La notebook se muestra difícil de leer. No se aprovechan los recursos que ofrece la plataforma.
2. APLICACIÓN DE CONOCIMIENTOS	El problema se resuelve en su totalidad y se aplican de manera óptima los conocimientos adquiridos en este módulo.	El problema se resuelve en su totalidad y se aplican de manera eficiente algunos de los conocimientos adquiridos en este módulo.	El problema se resuelve parcialmente o se aplican de manera correcta algunos de los conocimientos adquiridos en este módulo.	El problema no se resuelve o se aplican conocimientos básicos que no fueron adquiridos en este módulo.
<b>REPORTE</b>				
1. FORMATO	El reporte se apegue al formato establecido, está escrito sin faltas de ortografía y cuenta con una redacción clara.	El reporte se apegue al formato establecido, está escrito con algunas faltas de ortografía pero cuenta con una redacción clara.	El reporte se apegue al formato establecido, está escrito con algunas faltas de ortografía o cuenta con una redacción poco clara.	El reporte se apegue al formato establecido, está escrito con varias faltas de ortografía y cuenta con una redacción poco clara.
2. RESULTADOS	El reporte expone claramente los resultados y argumenta claramente su pertinencia y validez.	El reporte expone claramente los resultados y argumenta con cierta claridad su pertinencia y validez.	El reporte expone los resultados con cierta claridad o argumenta vagamente su pertinencia y validez.	El reporte expone los resultados con muy poca claridad o no argumenta su pertinencia y validez.

### Formato de la presentación y reporte

Tanto la presentación como el reporte deberán contar con la siguiente estructura:

1. Portada (título y nombre del/de la participante).
2. Introducción (planteamiento del problema).
3. Objetivos.
4. Metodología.
5. Resultados y discusión



## Acreditación del diplomado

Para acreditar el diplomado y obtener el diploma correspondiente, se debe:

1. Contar con un mínimo de asistencia del 80%.
2. Acreditar todos los módulos con calificación mínima de 7. Opcionalmente y para el alumno/a que lo solicite, se dará constancia de la calificación obtenida.

## Requisitos de conexión

- Las sesiones por la plataforma [zoom](#), por lo que cada participante debe asegurar una buena conexión a internet. Se recomienda contar con un servicio de al menos 20 o 30 MB por participante.
  - El acceso para los días **jueves** es el siguiente: <https://us02web.zoom.us/j/82381178076?pwd=bjV0Q0lFRW1WYi9wR2VORlFmcUZmZz09> Código de acceso: 829932.
  - El acceso para los días **sábado** es el siguiente: <https://us02web.zoom.us/j/88244447792?pwd=WjB3Zld5OGc3RmRJV2JWbU9hVCTvZz09> Código de acceso: 739755.
- Se llevará un registro de asistencia.
- Se **requiere una cuenta de Gmail** para poder acceder a los contenidos del curso.
- Una vez que se tenga una cuenta de Gmail, se sugiere explorar la plataforma de programación que será Google Colab: <https://colab.research.google.com/>
- El repositorio desde donde se cargará el material de los módulos es el siguiente: <https://github.com/DCDPUAEM/DCDP>
- Se creará un grupo de [classroom](#) para que los/las participantes puedan subir sus tareas y proyectos.

## Propiedad Intelectual

El material del diplomado como videos, presentaciones, imágenes, scripts o datos, están sujetos a derechos de autor y son propiedad intelectual de sus autores y fuentes dentro o fuera del Diplomado. En respeto a estos derechos, queda prohibido divulgar, publicar o compartir dicho material, en su totalidad o parcialmente, por medios electrónicos o físicos.

## Contenidos del Diplomado

Los Módulos han sido diseñados siguiendo pautas académicas con el propósito de impartir conocimiento y favorecer el proceso de enseñanza-aprendizaje. Sin embargo, se fomenta también la libertad de cátedra, por lo que cada instructor podrá modificar el contenido del Módulo a su cargo, si así lo juzga conveniente. Es responsabilidad de cada participante estudiar conceptos o temas previos a la materia de cada Módulo, que son requisito o base de conocimiento para alcanzar el mejor rendimiento posible y aprovechar al máximo los contenidos del diplomado.