

Nivel avanzado

Análisis de Datos y Aprendizaje Automático



Datos e
Inteligencia Artificial

SADOSKY
CAPACITACIONES



Análisis de datos y aprendizaje automático

Fundamentación

Actualmente, la demanda de profesionales con conocimientos en Ciencia de Datos (CD) e Inteligencia Artificial (IA) está en constante crecimiento, impulsada por la expansión de estas tecnologías en diversos sectores como finanzas, biología y salud. Desde asistentes virtuales y sistemas de recomendación hasta modelos de predicción y generación de contenidos, su aplicación es cada vez más amplia.

Sin embargo, el desarrollo e implementación de sistemas de CD e IA requieren conocimientos especializados. Aplicarlos sin una comprensión adecuada puede generar riesgos significativos, como interpretaciones erróneas de los resultados o errores en la toma de decisiones, afectando tanto a quienes desarrollan estos sistemas como a sus usuarios, ya sea de forma directa o indirecta.

Para responder a esta necesidad, desde el año 2023 desde la Fundación Sadosky ofrecemos una Capacitación Avanzada en CD e IA, dirigida a profesionales que tengan conocimientos de programación y deseen iniciarse en el camino de la ciencia de datos y la IA.

Objetivo de la capacitación

Desarrollar competencias fundamentales en Ciencia de Datos e Inteligencia Artificial que permitan su implementación efectiva y responsable en distintos contextos de aplicación.

Contenidos

Módulo 1: Fundamentos en ciencia de datos e inteligencia artificial

- Conceptos y metodología
- Reforzando programación en Python

Módulo 2: Análisis exploratorio de datos

- Bases del Análisis Exploratorio de Datos
- Explorando Pandas: Manipulación de datos
- Estadística descriptiva



Módulo 3: Preprocesamiento de datos

- Limpieza y preparación de datos
- Valores faltantes: definición y riesgos
- Detección de valores faltantes: técnicas de detección y visualización
- Tratamiento de valores faltantes: métodos de Imputación

Módulo 4: Visualización de datos

- Fundamentos de visualización de datos
- Herramientas de visualización
- Narrativa con datos

Módulo 5: Modelos de clasificación

- Introducción al aprendizaje automático
- Modelos de clasificación
- Métricas para evaluar modelos de clasificación
- Herramientas: Scikit-learn

Módulo 6: Modelos de regresión

- Regresión lineal y mínimos cuadrados
- Funciones de costo
- Optimización y método de descenso del gradiente
- Diagnóstico y métricas de evaluación
- Regresión no lineal y otros modelos

Módulo 7: Aprendizaje no supervisado

- Aprendizaje no supervisado
- Clustering de datos
- k-medias (K-means)
- Encontrar el número óptimo de clústeres
- La maldición de la dimensión
- Reducción de la dimensionalidad

Módulo 8: Introducción a las redes neuronales

- Introducción a las redes neuronales
- Perceptrón simple y multicapa
- Funciones de activación
- Entrenamiento de una red neuronal

Organización y cronograma de cursada

La capacitación se organiza en **ocho módulos** a ser cursados en **cuatro meses**, y diseñados para fortalecer progresivamente los conceptos, metodologías, técnicas y herramientas esenciales en proyectos de CD e IA.



Cada módulo incluye materiales teóricos, recursos audiovisuales, notebooks en Python y cuestionarios de evaluación. Los contenidos están pensados para ser recorridos de manera flexible, permitiendo a los cursantes avanzar a su propio ritmo y según sus intereses dentro de los tiempos sugeridos. Además, se proponen ejercicios opcionales y lecturas complementarias para profundizar en los temas abordados.

Además, se pondrán a disposición **ocho encuentros virtuales sincrónicos** de aproximadamente dos horas y con una frecuencia quincenal, donde podrán interactuar con la docente y sus pares para resolver dudas y afianzar conceptos. Las fechas de estos encuentros serán comunicadas a través de la plataforma de capacitación.

Adicionalmente, los cursantes podrán optar por la realización de un **proyecto integrador**. Para esto, se han planificado, a lo largo de la cursada, entregas prácticas que corresponden a avances parciales de este proyecto, permitiendo aplicar los conocimientos adquiridos en un caso concreto con datos.

Al finalizar los módulos de estudio y las entregas prácticas, se accederá a una instancia final que incluye una evaluación integradora y el cierre de la formación.

**Detalle de cronograma tentativo**

MÓDULO	SEMANA	FECHA
1	1 y 2	Del 18 de mayo al 2 de junio
Espacio de tutoría sincrónica. 4 de junio.		
2	3 y 4	Del 3 al 16 de junio
Espacio de tutoría sincrónica. 18 de junio.		
3	5 y 6	Del 17 al 30 de junio
Espacio de tutoría sincrónica. 2 de julio.		
4	7 y 8	Del 1 al 14 de julio
Espacio de tutoría sincrónica. 16 de julio. Por grupos, enfocada en la entrega de la primera parte del proyecto integrador.		
5	9 y 10	Del 15 al 28 de julio
Espacio de tutoría sincrónica. 30 de julio		
6	11 y 12	Del 29 de julio al 11 de agosto.
Espacio de tutoría sincrónica. 13 de agosto. Por grupos, enfocada en la entrega de la segunda parte del proyecto.		
7	13 y 14	Del 12 al 25 de agosto
Espacio de tutoría sincrónica. 27 de agosto.		
8 Proyecto integrador	15 y 16	Del 26 de agosto al 8 de septiembre.
Presentación de proyectos finales. 10 de septiembre.		
30 días de corrección por parte del/la docente		

Los detalles de asistencia y aprobación de la capacitación serán detallados en plataforma.



Evaluación y cierre de la cursada

Se puede obtener la certificación de la capacitación de dos maneras que exigen distintos requisitos:

- **Certificado de aprobación. Requisitos:**
 - Aprobar las actividades obligatorias intermedias, es decir, las de cada módulo.
 - Aprobar el cuestionario final.
- **Certificado de aprobación con evaluación. Requisitos:**
 - Aprobar las actividades obligatorias intermedias.
 - Aprobar el cuestionario final.
 - Presentar y aprobar el Proyecto Integrador.

Sobre el proyecto integrador (PI)

Para consolidar los conocimientos adquiridos, la formación propone, de manera opcional, la realización de un **proyecto integrador** en el que se podrán aplicar técnicas y metodologías de Ciencia de Datos e Inteligencia Artificial en un caso concreto. Este proyecto se desarrollará en **dos entregas parciales** obligatorias y una **entrega final** en formato notebook, que se presentará oralmente en un encuentro sincrónico con la docente y los colegas que hayan cursado en la cohorte.

Su objetivo es ofrecer una experiencia práctica integral que abarque desde el análisis exploratorio de datos hasta la construcción de modelos predictivos, promoviendo asimismo el desarrollo de habilidades técnicas, analíticas y comunicativas.

Para acceder a la evaluación del proyecto integrador, se requiere: aprobar los cuestionarios de cada módulo, completar las entregas parciales en tiempo y forma, y superar un test teórico final.



Criterios de evaluación del PI

CRITERIOS	REGULAR	BIEN	MUY BIEN	DESTACADO
Proceso de aprendizaje y actitud ante el feedback	Participación irregular. No incorpora devoluciones ni evidencia mejora.	Cumple con las actividades y aplica parcialmente las sugerencias recibidas.	Integra el feedback de manera consistente y evidencia evolución en su desempeño.	Reflexiona críticamente sobre su proceso, integra estratégicamente las devoluciones y demuestra autonomía y mejora significativa.
Dominio conceptual	Comprensión superficial o fragmentada de los conceptos centrales.	Aplica correctamente los conceptos fundamentales del curso.	Integra conceptos con coherencia y fundamenta decisiones técnicas.	Articula críticamente el marco conceptual, establece relaciones complejas y aporta perspectiva propia.
Terminología disciplinar	Uso impreciso, limitado o coloquial del vocabulario específico.	Emplea terminología disciplinar básica de manera adecuada.	Utiliza vocabulario técnico amplio con precisión y pertinencia.	Domina laRoboto Condensed terminología especializada con solvencia profesional e integra aportes interdisciplinarios.
Desarrollo técnico y metodológico del proyecto integrador	Proyecto incompleto o con debilidades técnicas y/o metodológicas significativas.	Cumple con los requisitos técnicos y metodológicos establecidos.	Presenta solidez técnica, coherencia metodológica y aportes pertinentes.	Integra conocimientos con profundidad, rigor metodológico y calidad profesional, evidenciando proyección estratégica.
Comunicación y presentación profesional	Presentación desorganizada o poco clara.	Comunicación clara y estructurada acorde a los estándares requeridos.	Manejo sólido del discurso técnico y de los recursos expresivos.	Presentación profesional, persuasiva y con identidad propia; posiciona estratégicamente su trabajo.



Bibliografía

- Huyen, C. (2022). *Designing Machine Learning Systems: An Iterative Process for Production-Ready Applications*. O'Reilly Media.
- James, G., Witten, D., Hastie, T. (2023). *An Introduction to Statistical Learning with Applications in Python*. Springer.
- Mukhiya, SK. y Ahmed U. (2020). *Hands-On Exploratory Data Analysis with Python: Perform EDA techniques to understand, summarize, and investigate your data*. Packt Publishing Ltd.
- Müller, A. y Guido, S. (2016). *Introduction to Machine Learning with Python*. O'Reilly.
- Zervou, M. (2024). *Python Data Cleaning and Preparation Best Practices: A practical guide to organizing and handling data from various sources and formats using Python*. Packt Publishing Ltd.

Créditos

Equipo de producción de Sadosky Capacitaciones para este curso:

- Coordinación: Julián Dabbah
- Contenidos: PCDIA - Fundación Sadosky
- Diseño gráfico y producción audiovisual: Fabio Viale, Facundo Manini, Federico Rey
- Didactización y maquetación: Débora Cingolani
- Corrección: Florencia Acher