



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

ANEXO I - PROTOCOLO XIII

a) Denominación.....	1
b) Destinatarios.....	1
c) Requisitos de ingreso.....	1
d) Objetivos.....	1
e) Justificación.....	2
f) Pertinencia respecto a la unidad académica.....	3
g) Estructura.....	4
h) Contenidos de cada unidad o módulo.....	5
i) Modalidad de cursado.....	7
j) Cronograma de dictado y carga horaria total expresada en horas y créditos.....	7
k) Nómina de equipo directivo y de docentes y CV nominal de cada uno.....	8
l) Modalidades de evaluación.....	9
m) Requisitos de aprobación.....	11
n) Bibliografía.....	11
o) Modelo de Certificado a otorgar.....	12



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

a) Denominación

Diplomatura Universitaria de Formación Continua en Data Science & Machine Learning

b) Destinatarios

La Diplomatura estará abierta al público en general, con especial énfasis en estudiantes de grado, graduados, investigadores, y profesionales de cualquier rubro que deseen aprovechar el uso de grandes volúmenes de datos. Está diseñada para quienes buscan aprender a analizar datos de manera eficiente y desarrollar modelos de aprendizaje automático (Machine Learning), con el fin de implementar técnicas prescriptivas y aplicar herramientas modernas que optimicen y potencien su desempeño en diversas áreas de trabajo.

c) Requisitos de ingreso

Requisitos académicos

1. Se recomienda conocimientos previos en programación orientada a objetos, no excluyente.
2. Se recomienda conocimientos en estadística básica, no excluyente.

Requisitos administrativos:

1. Computadora con conexión a internet y micrófono.
2. Abonar el arancel.

d) Objetivos

Objetivo General

Formar profesionales competentes en el manejo y análisis de datos complejos, capaces de desarrollar e implementar soluciones innovadoras a través de modelos de Machine Learning.

Objetivos Específicos



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

1. Desarrollar la capacidad de manipular, explorar, limpiar y preparar datos provenientes de diversas fuentes, asegurando su calidad y aplicando técnicas avanzadas de Machine Learning para crear modelos predictivos.
2. Aplicar el ciclo de trabajo de la Ciencia de Datos, novedoso pipeline de trabajo empleado en la industria, para desarrollar soluciones basadas en datos en entornos organizacionales públicos o privados.
3. Crear visualizaciones de datos efectivas y comunicar los hallazgos de manera clara,
4. Fomentar la participación en comunidades científicas y proyectos colaborativos, aplicando las competencias adquiridas para desempeñar el rol de Científico de Datos.

e) Justificación

La Ciencia de Datos es el estudio de la materia prima más abundante en el universo digital: los datos. Se trata de extraer, analizar, visualizar, gestionar y almacenar datos, para ayudar a las empresas a tomar decisiones basadas en datos, a desarrollar sistemas inteligentes para generar nuevas aplicaciones y potenciar sus operaciones tradicionales.

Específicamente, se identifican las siguientes necesidades en ámbitos industriales y académicos:

1. Demanda de perfiles Data Scientist Trainee y Jr, capaces de trabajar dentro de ambientes productivos y aportar al desarrollo de soluciones de datos integrales. El BLS Employment Projections proyecta un crecimiento de 35 puntos porcentuales en la demanda a nivel mundial en el período 2022-2032.
2. Demanda de profesionales con conocimientos en lenguajes de programación, entre los cuales Python destaca por ser el lenguaje de propósito general más utilizado en Argentina, según los resultados aportados por la última encuesta de SysArmy.



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

3. Excelente oportunidad de desarrollo profesional: la U.S. News & World Report, posiciona al Data Scientist como el cuarto mejor trabajo en tecnología, entre decenas de opciones.
4. El Machine Learning, como elemento central del proceso de montaje, selección y deploy de modelos, emerge como la habilidad más querida entre las empresas de datos.

La Diplomatura Universitaria de Formación Continua en Data Science & Machine Learning tiene como propósito formar profesionales capaces de enfrentar desafíos comerciales complejos mediante la aplicación de técnicas avanzadas de ciencia de datos y modelos de aprendizaje automático (Machine Learning). Este programa está diseñado para introducir a los estudiantes en el rol de Científico de Datos Trainee, complementando perfiles de analistas, investigadores y otros profesionales que buscan profundizar en el manejo y análisis de datos. Además, busca contribuir al desarrollo tecnológico del país, cubriendo la creciente demanda de profesionales en el ámbito de la ciencia de datos y estableciendo un puente entre universidades, empresas y alumnos.

f) Pertinencia respecto a la unidad académica

El avance tecnológico y la superabundancia de datos han producido un aumento marcado en la necesidad de contar con talento analítico en las empresas. Es en función de esta necesidad que se plantea la Diplomatura Universitaria de Formación Continua en Data Science & Machine Learning, como una oportunidad para aportar a la misión de la FCEFYN de nutrir la demanda de profesionales en áreas estratégicas. Específicamente, la Diplomatura tiene una vinculación directa con carreras de la FCEFYN: Ingeniería en Computación, Ingeniería Electrónica e Ingeniería Biomédica. Así también carreras con fuerte componente científico como Biología y Geología. La Ciencia de Datos en particular, se ha transformado en los últimos años en una salida laboral importante para los graduados de esas carreras. Específicamente, el programa académico presenta notables oportunidades de aprendizaje para estudiantes y graduados de Ingeniería Biomédica al presentar contenidos que permiten iniciarse en el desarrollo de modelos de aprendizaje automático con aplicaciones directas en el campo de la medicina: procesamiento de imágenes para detección automática de patologías, procesamiento de historias clínicas, automatización de triaje y evaluación de riesgo, etc.



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

Asimismo, la Diplomatura incluye contenidos de gestión de proyectos, enfoque sistémico, programación orientada a objetos y manejo de herramientas tecnológicas de vanguardia, conjunto de aptitudes centrales en la enseñanza de la Ingeniería y las Ciencias en el mundo actual.

Por último, si bien dentro del ámbito de la Facultad se pueden encontrar formaciones previas dentro del área del Data Science, la Diplomatura se diferencia en que ofrece el desarrollo de un toolkit (conjunto de herramientas) integral para desempeñarse como Data Scientist Trainee, con foco en el Machine Learning.

g) Estructura

La Diplomatura se compone de 8 módulos secuenciales, a dictarse a distancia. Se seguirá una metodología de aprendizaje por proyectos en la que se acompañará al alumno, con el docente de clase y un tutor de la cursada, a crear soluciones innovadoras con datos reales. El programa se completa con 3 sprints y finaliza con la entrega de un proyecto en el que se ponen en práctica los conceptos aprendidos. El objetivo es que el alumno desarrolle su portfolio personal. En resumen, la propuesta aborda 3 entregas desarrolladas durante 6 meses, con dos clases por semana de 2 horas cada clase.

Asimismo, se contará con material de práctica semanal, que será provisto mediante Jupyter Notebooks y/o Colab Notebooks y abordará las temáticas vistas en las clases sincrónicas, ampliando los conceptos teóricos y reforzando la ejercitación práctica. Las Notebooks se entregarán en versiones a completar y luego de un período prudente la versión corregida con un modelo de resolución propuesta. Se contará con espacios asincrónicos y/o sincrónicos para que los estudiantes aborden sus resoluciones personales a las problemáticas presentadas en las Notebooks que servirán como un puntapié para la resolución de los sprints.

Grilla de metodología con carga horaria

Módulo/Metodología	Clases sincrónicas teóricas-prácticas (horas)	Trabajos Prácticos (tiempo personal, instancias de tutorías)
--------------------	---	--



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

Módulo I	20	25
Módulo II	8	10
Módulo III	16	20
Módulo IV	8	10
Módulo V	20	25
Módulo VI	16	20
Módulo VII	8	10
Módulo VIII	4	5
Sprints y trabajo integrador	0	25
Total de horas	100	150

h) Contenidos de cada unidad o módulo

MÓDULO 1: Introducción Data Science

Objetivo: Introducir a la dinámica de trabajo de un data scientist y aprender a usar las librerías Numpy y Pandas para manipulación, limpieza y cálculo con datos

- Introducción: qué es un dataset, tipos de datos, tipos de problemas.
- Introducción Data Science: workflow.
- Introducción a programación básica.
- Presentación Jupyter Notebooks y bibliotecas NumPy, Pandas, Matplotlib.
- Breve introducción NumPy: tipos de datos y operaciones.
- Exploración de datos: Pandas.

MÓDULO 2: Feature Engineering

Objetivo: Profundizar las prácticas de limpieza y manipulación de datos. Se introducirán buenas prácticas de análisis exploratorio de datos

- Feature engineering conceptos: outliers, missings, categoricas, dummies, nuevas variables.
- Feature engineering Pandas: outliers, missings, dummies, nuevas variables.



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

- Feature engineering Scikit Learn: outliers, missings, categoricals.

MÓDULO 3: Machine Learning I. Introducción algoritmos de aprendizaje supervisado

Objetivo: Con las bases de trabajo en Python más consolidadas, se introduce a la metodología de trabajo de Machine Learning, junto a una de sus librerías más importantes, Scikit-Learn

- Evaluación de modelos: training/testing, matriz de confusión.
- Algoritmo KNN.
- Algoritmo Decision trees.
- Clasificación y regresión con estos algoritmos.
- Overfitting/Underfitting.
- Cross Validation.
- Pipelines en scikit-learn.

MÓDULO 4: Selección de modelos

Objetivo: Hacer foco en la correcta evaluación de distintos modelos para seleccionar el más adecuado para nuestra problemática

- Tradeoff bias/variance.
- Optimización de parámetros: GridSearch y hyperopt.
- Selección stepwise de variables: forward, backward.
- Otras métricas: AUC, F1, kappa.

MÓDULO 5: Machine Learning II. Procesamiento de texto y Algoritmos de aprendizaje supervisado avanzados.

Objetivo: Abordar problemas más avanzados como text mining, redes neuronales, modelos de ensamble bagging y boosting.

- Introducción a procesamiento de texto: bag of words, tf-idf, tokens, stemming.
- SVM.
- Ensamblados básicos: voting, promedios.
- Ensamblados bagging: random forests.
- Ensamblados boosting: adaboost, xgboost, catboost.
- Redes Neuronales: perceptrón y backpropagation.
- Keras: Redes Neuronales multicapas.



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

MÓDULO 6: Machine Learning III. Algoritmos de aprendizaje no supervisado y sistemas de recomendación.

Objetivo: Profundizar sobre los modelos no supervisados y sus casos de uso más frecuentes, así como introducir en la construcción de algoritmos de recomendación y sus técnicas.

- Introducción aprendizaje no supervisado.
- Clustering: kmedias, dbscan.
- Reducción de dimensionalidad: PCA.
- Introducción sistemas de recomendación.
- Reducción de dimensionalidad: SVD.
- Sistemas de recomendación Filtro colaborativo.
- Evaluación de sistemas de recomendación.

MÓDULO 7: Deploy de modelos

Objetivo: Disponibilizar los modelos y trabajar con APIs

- Persistencia de modelos scikit-learn/python.
- Desarrollo API para acceder al modelo y predicciones.
- Puesta en producción Cloud.

MÓDULO 8: Cómo seguir

Objetivo: Dar los lineamientos finales e integrar los conocimientos adquiridos y dar un pantallazo de las nuevas tendencias en Machine Learning

- Herramientas cloud: gcp/aws/azure.
- Machine learning on demand: floyhub, paperspace, crestle.
- Kaggle y cómo perfeccionarse.

i) Modalidad de cursado

Modalidad presencial con encuentros sincrónicos mediados con tecnología.

- Plataforma de videoconferencia a definir (Sugerido: Google Meet/Zoom)
- Acceso a través de plataforma de foro y mensajería (Sugerido: Google Classroom/Slack).



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

j) Cronograma de dictado y carga horaria total expresada en horas y créditos

Carga horaria

La diplomatura se dictará durante 25 semanas, con un total de 250 horas divididas en 100 horas de clase sincrónica, distribuidas en 50 encuentros virtuales, más 150 horas de práctica asíncrona por parte del estudiante. Estas últimas incluirán la realización de actividades prácticas, de investigación y de laboratorio. A tales fines, existirán canales de comunicación dispuestos para funcionar fuera del horario de clase.

Cronograma

Table with 2 parts showing module distribution over 25 weeks. Part 1 (weeks 1-13) includes modules like 'Introducción Data Science', 'Feature Engineering', 'Introducción algoritmos de aprendizaje supervisado', and 'Selección de modelos'. Part 2 (weeks 14-25) includes 'Procesamiento de texto y Algoritmos de aprendizaje supervisado avanzados', 'Algoritmos de aprendizaje no supervisado y sistemas de recomendación', 'Deploy de modelos', and 'Cómo seguir'.

Créditos CRE

Corresponde a un total de 10 créditos CRE.

k) Nómina de equipo directivo y de docentes y CV nominal de cada uno

Table with 4 columns: Rol, Docente UNC, Nombre y Apellido, Correo electrónico. Row 1: Director - Docente UNC, Prof. Titular Modelos y Simulación -, Mgtr. Ing. Alejandro, alejandro.ambrosini@unc.edu.ar



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

	FCEfyN - UNC	Ambrosini	
Coordinador operativo	-	Mgtr. Ing. Martin Jaureguy	-
Capacitador	-	Mgtr. Ing. Diego Rojas	-
Capacitador	-	Ing. Medina Damián Prámparo	-
Capacitador	-	MBA Tec. Julieta Kogan Moyano	-
Capacitador	-	Lic. Facundo Buffa	-

CV con libre acceso en el siguiente link. Serán presentados al expediente como Anexos.

https://drive.google.com/drive/folders/17ZGgON9GN5Z4DwDPSa0myDlsil6m_Jtj?usp=sharing

I) Modalidades de evaluación

La evaluación se hará sobre cada entrega del sprint por parte del equipo evaluador compuesto por el capacitador y el mentor a cargo de la comisión.

SPRINT I

Objetivos:

Crear familiaridad con el abordaje real de un problema basado en datos. Comienza con una etapa de EDA (Análisis Exploratorio de Datos), primer elemento de la cadena de un proceso de Data Science.

Se emplea el lenguaje Python y librerías de procesamiento, limpieza y visualización de datos. Además, se emplearán los primeros modelos de Machine Learning.

Preguntas para despertar la creatividad de los estudiantes

- Si tengo un problema, ¿qué datos necesito para poder llegar a una posible solución?
- Los datos que requiero, ¿tengo acceso a ellos, hace falta recolectarlos, cómo los puedo capturar?



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

- Si ya dispongo de mi set de datos, ¿qué preguntas me permite responder?

Proyecto:

Machine Learning desde cero: mi primer modelo

Se provee una problemática donde el estudiante deberá armar el caso de uso, definir el/los objetivos e identificar si cuenta con los datos requeridos para esbozar la solución, también deberá realizar un análisis exploratorio del set de datos y utilizar un modelo simple de Machine Learning como solución propuesta.

SPRINT II

Objetivos:

Profundizar en la estadística y matemática que está detrás de cada modelo, dando un salto en complejidad en la comprensión de Machine Learning.

Hacer foco en la interpretación de resultados y la correcta evaluación y comparación de los mismos.

Entender el comportamiento de algoritmos de Machine Learning y distinguir predicciones para su uso crítico.

Proyecto:

Advanced Machine Learning: potenciando mi solución

Se presenta un desafío consistente en aplicar técnicas de ingeniería de features enriqueciendo la información disponible. El estudiante deberá utilizar y desarrollar algoritmos avanzados de Machine Learning e interpretar los resultados.

Preguntas para despertar la creatividad de los estudiantes

- ¿Se puede mejorar aún más los resultados?
- ¿Cuáles serían tus conclusiones?
- ¿Por qué se seleccionó este modelo sobre otros

SPRINT III

Objetivos:

Presentar al estudiante mecanismos y metodologías para resolver problemas de dominios específicos según su interés particular.

Introducir en aplicaciones empresariales y científicas de los modelos.

Presentar casos de uso como NLP, Modelos de Lenguaje, Recommender System, Detección de Imágenes y Times Series.

Estimular la autonomía para el desarrollo de soluciones.



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

Proyecto

Machine Learning Solutions: desarrollando aplicaciones modernas

Se presenta un proyecto que deberá ser resuelto implementando modelos novedosos, donde el Machine Learning ha marcado un antes y un después. El estudiante deberá utilizar herramientas de punta para presentar sus resultados.

Se estimula al estudiante en la presentación de resultados, disponibilización de modelos y storytelling, mostrando la importancia de este apartado: No basta con tener la mejor solución, ni haber realizado un trabajo brillante sino es disponibilizado y comunicado de manera efectiva con sus respectivas justificaciones.

Se considerará

- Completitud de cada SPRINT
- Participación en clases
- Verificación de la transferencia de los contenidos desarrollados
- Participación asincrónica en plataformas
- Resolución de Notebooks semanales

m) Requisitos de aprobación

Nota final

El estudiante deberá aprobar con nota igual o superior a 6 puntos sobre 10 cada sprints, lo cual indica un cumplimiento del 60% de cada una de las entregas.

Asistencia

Asistencia al 75% de las clases sincrónicas. Se podrá eximir al estudiante de la asistencia bajo expreso pedido del mismo, en casos excepcionales. Los alumnos que aprueben las exigencias descritas recibirán un certificado de aprobación.

n) Bibliografía



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

1. de la Torre, J. (2023). Redes Generativas Adversarias (GAN): Fundamentos Teóricos y Aplicaciones.
2. Vaca Serrano, A., Garcia Subies, G., Montoro Zamorano, H., Aldama Garcia, N., Samy, D., Betancur Sanchez, D., ... & Barbero Jimenez, A. (2022). RigoBERTa: A State-of-the-Art Language Model For Spanish.
3. Cañete, J., Donoso, S., Bravo-Marquez, F., Carvallo, A., & Araujo, V. (2022). ALBETO and DistilBETO: Lightweight Spanish Language Models.
4. Cutiérrez-Fandiño, A., Armengol-Estapé, J., Pàmies, M., Llop-Palao, J., Silveira-Ocampo, J., Carrino, C. P., ... & Villegas, M. (2021). MarIA: Spanish Language Models.
5. Lakshmanan, V., Robinson, S., & Munn, M. (2020). Machine Learning Design Patterns: Solutions to Common Challenges in Data Preparation, Model Building, and MLOps. O'Reilly Media.
6. Géron, A. (2019). Aprende Machine Learning con Scikit-Learn, Keras y TensorFlow: Conceptos, herramientas y técnicas para conseguir sistemas inteligentes (2ª ed.). O'Reilly Media.
7. Burkov, A. (2019). El libro de Machine Learning de 100 páginas. Andriy Burkov.
8. Raschka, S., & Mirjalili, V. (2019). Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2. Packt Publishing.
9. Chollet, F. (2018). Deep Learning con Python. Manning Publications.
10. Zheng, A., & Casari, A. (2018). Feature Engineering for Machine Learning. O'Reilly Media.
11. McKinney, W. (2017). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython (2ª ed.). O'Reilly Media.

o) Modelo de Certificado a otorgar



ACTA ACUERDO – PROTOCOLO XIII – ANEXO CONVENIO

RD 2192/2019

ICARO



<<CÓDIGO>>

UNC FCEFyN

Diplomatura Universitaria de Formación Continua en Data Science & Machine Learning

DÍA / MES / AÑO

Certificamos que <<Nombre>> <<Apellido>>, DNI: <<DNI>>, ha realizado y aprobado la **Diplomatura Universitaria de Formación Continua en Data Science & Machine Learning**, dictada durante el año 2025 por ICARO Asociación Civil en conjunto con la Secretaría de Extensión de la **Facultad de Ciencias Exactas Físicas y Naturales** de la **Universidad Nacional de Córdoba**.

Carga horaria: 250 horas
CRE: 10
Resolución n°:

La presente Diplomatura no constituye una carrera universitaria y por lo tanto no otorga título habilitante para el ejercicio profesional (RHCS N° 64/2024).


ING. MTR. PABLO RICCARDI
SECRETARIO
FACULTAD DE CIENCIAS EXACTAS FÍSICAS Y NATURALES
UNIVERSIDAD NACIONAL DE CÓRDOBA


ING. LUIS BOSCH
SECRETARIO DE EXTENSIÓN
FACULTAD DE CIENCIAS EXACTAS FÍSICAS Y NATURALES
UNIVERSIDAD NACIONAL DE CÓRDOBA


ING. SANTIAGO ALCALDE
APODERADO DE ICARO
ASOCIACIÓN CIVIL