



Universidad  
Nacional  
de Córdoba



Facultad de Matemática,  
Astronomía, Física y  
Computación

EXP-UNC: 60737/2017

## VISTO

La propuesta para el dictado de la "Diplomatura en Ciencia de Datos, Aprendizaje Automático y sus Aplicaciones" presentada por la Dra. Laura ALONSO ALEMANY y los Dres. Jorge Adrián SÁNCHEZ y Oscar REULA; y

## CONSIDERANDO

Que la propuesta del plan de estudios ha sido analizada por la SECRETARIA DE ASUNTOS ACADEMICOS de la UNIVERSIDAD NACIONAL DE CÓRDOBA y la misma no ha formulado objeciones.

Que se han realizado consultas a representantes del sector empresario relacionado al área de la informática y Ciencias de la Computación de Córdoba y se han analizado las demandas del mismo.

Que esta Diplomatura tiene como objetivo cubrir las necesidades detectadas en la formación de recursos humanos del sector socio-productivo y académico de la región en este área de conocimiento.

Que la propuesta presentada se enmarca en la Resolución HCS 483/2015 y será autofinanciada.

**Por ello,**

EL CONSEJO DIRECTIVO  
DE LA FACULTAD DE MATEMÁTICA, ASTRONOMÍA, FÍSICA Y COMPUTACIÓN

RESUELVE:

ARTÍCULO 1º: Aprobar el dictado de la "Diplomatura en Ciencia de Datos, Aprendizaje Automático y sus Aplicaciones" según se detalla en el Anexo que forma parte de la presente Resolución.

ARTÍCULO 2º: Establecer que el funcionamiento de esta Diplomatura estará vinculado, académicamente a la Secretaría Académica y administrativamente a la Oficina de Vinculación Tecnológica de esta Facultad.

ARTÍCULO 3º: Crear la figura de Coordinador de la Diplomatura en Ciencia de Datos, Aprendizaje Automático y sus Aplicaciones.

ARTÍCULO 4º: Establecer como funciones del Coordinador las siguientes:

a) Representar a la diplomatura ante instituciones públicas y privadas, cuando corresponda, sin desmedro de las atribuciones de el/la Secretario/a Académico/a y con el acuerdo de éste/a.

EXP-UNC: 60737/2017


b) Colaborar con la Secretaría Académica y la Oficina de Vinculación Tecnológica en todas las actuaciones que atañen al funcionamiento de la diplomatura.

c) Interactuar con el Consejo Académico Profesional de la Diplomatura y el plantel docente a fin de asesorarse de las necesidades de la misma.

ARTÍCULO 5º: Establecer que la Diplomatura contará con un Comité Académico Profesional (CAP), el cual estará presidido por el Coordinador.

ARTÍCULO 6º: Establecer que al igual que las Maestrías y Especializaciones que se dictan en la Facultad, el Consejo Directivo deberá establecer el monto del curso, a propuesta del Coordinador de la Diplomatura.

ARTÍCULO 7º: Notifíquese, publíquese y archívese.

 DADA EN LA SALA DE SESIONES DEL CONSEJO DIRECTIVO DE LA FACULTAD DE MATEMÁTICA, ASTRONOMÍA, FÍSICA Y COMPUTACIÓN A VEINTISÉIS DÍAS DEL MES DE FEBRERO DE DOS MIL DIECIOCHO.

**RESOLUCIÓN CD N° 14/2018**



Dra. SILVIA PATRICIA SILVETTI  
SECRETARIA GENERAL  
FaMAF

  
Dra. Ing. MIRTA IRIONDO  
DECANA  
FaMAF

EXP-UNC: 60737/2017

## ANEXO

### 1. DATOS DEL CURSO:

a. Nombre de la diplomatura

**Diplomatura en Ciencia de Datos, Aprendizaje Automático y sus Aplicaciones**

b. Unidad académica

FaMAF

c. Período lectivo

Abril a Noviembre

d. Cantidad de RTF (Reconocimiento de Trayectos Formativos)

6 RTF (180 horas, 120 presenciales y 60 virtuales)

e. Horario de clases presenciales

Fines de semana alternos, viernes alterno de 17 a 21 y sábado de 10 a 14

f. Lugar en que se desarrollarán las clases

FaMAF

g. Número mínimo y máximo de estudiantes:

Mínimo 40 máximo 50

h. Perfil de los estudiantes que pueden asistir al curso

El curso se dirige principalmente a estudiantes y profesionales de la computación, ingenierías o de actividades que involucren el trabajo y modelado estadístico de datos.

### 2. DATOS DEL CUERPO ACADÉMICO:

El cuerpo académico constará de docentes de las universidades y de la industria. Se presentará oportunamente en un anexo. El coordinador será Jorge A. Sánchez, del área de Ciencias de la Computación ([jsanchez@famaf.unc.edu.ar](mailto:jsanchez@famaf.unc.edu.ar)).

EXP-UNC: 60737/2017

## PROGRAMA DEL CURSO O DIPLOMATURA

### FUNDAMENTACIÓN

En los últimos años se está desarrollando el área de ciencia de datos con ramificaciones en aprendizaje automático, inteligencia artificial y múltiples aplicaciones industriales. Este gran desarrollo ha generado una gran demanda en todos los ámbitos socio-productivos y académicos, y a la vez ha impulsado avances cada vez más rápidos. Esta demanda solo puede ser satisfecha si se reúnen esfuerzos tanto del sector académico como del sector socio-productivo.

En la ciencia de datos los **fundamentos teóricos** resultan imprescindibles para la práctica diaria y para mantenerse actualizado. Al mismo tiempo, resulta ineludible basarse en una rigurosa **metodología empírica**, basada en experimentación y evaluación. Es por esta razón que proponemos una diplomatura que proveerá el **vínculo** entre fundamentos teóricos y práctica para el análisis de problemas y diseño de soluciones basadas en el procesamiento de datos.

Los estudiantes desarrollarán madurez en el área a partir del estudio de casos en ejemplos concretos, aprendiendo al mismo tiempo las herramientas prácticas que les permitirán implementar y evaluar soluciones.

### OBJETIVOS

Proveer a los estudiantes el conocimiento y las herramientas para entender y aplicar técnicas analíticas apropiadas para la transformación y el análisis de datos.

Emplear con eficiencia técnicas estadísticas y de aprendizaje automático para análisis de datos. Conocer, diseñar e implementar soluciones basadas en aprendizaje automático.

Desarrollar habilidades para abordar efectivamente tareas que involucren grandes conjuntos de datos.

Evaluar y comunicar efectivamente los resultados del tratamiento de datos, interpretación, alcances e implicaciones de los resultados obtenidos.

### CONTENIDOS

#### 1. Materias Obligatorias

##### 1.1. Análisis y visualización de datos

Conceptos básicos (frecuentista). Probabilidad condicional. Regla de Bayes. Independencia.



EXP-UNC: 60737/2017

Independencia condicional. Ley de los grandes números. Estimadores de máxima verosimilitud. Estimadores puntuales y por intervalos. Testeo de hipótesis. Significancia. Investigación reproducible.

Conceptos sobre visión y percepción. Cómo utilizar gráficos para presentar información cuantitativa (tablas de contingencia, gráficos de barras, histogramas, formas de representación de datos categóricos y continuos). Visualización de datos multivariados. Animaciones, sonido y otras modalidades para presentar información.

### 1.2. Análisis exploratorio y curación de datos

Análisis exploratorio de datos. Ingesta de datos y filtrado de conjuntos de datos para ML. Generación de representaciones y diseño de experimentos. Exploración de datos y cómputo de sus estadísticas básicas. Generación de muestras según distribuciones.

Clustering y reglas de asociación. K-means y Mixturas Gaussianas. Reducción de dimensionalidad. PCA.

Detección de anomalías. Detección de outliers. Detección de ruido, errores, valores faltantes. Normalización.

### 1.3. Introducción al aprendizaje automático

Regresión. Intuición geométrica. Derivación por máxima verosimilitud. Regularización. Clasificación. Intuición geométrica. Regresión logística. Nociones de optimización convexa. Aprendizaje por descenso de gradiente. Clasificadores de margen máximo. Naïve Bayes. Modelos no paramétricos: vecinos más cercanos. Clasificación basada en reglas. Árboles de decisión. Aplicaciones. Problemas multiclase: one-vs-all, one-vs-one. Laboratorio(s).

Métricas y figuras de performance. Validación cruzada. Técnicas de reducción de varianza. Interpretación de resultados. Feature selection.

### 1.4. Aprendizaje supervisado

Métodos no lineales. Métodos basados en kernels. SVMs. Ingeniería de Kernels. Redes neuronales. Arquitecturas feed-forward. Backpropagation. Boosting y métodos ensemble.

Aprendizaje profundo. Redes convolucionales. Redes recurrentes. Otras arquitecturas. Datos secuenciales. Cadenas de Markov. HMM. Algoritmo Expectation-Maximization. Inferencia. Problemas. CRF. Aplicaciones a problemas de lenguaje natural y visión por computadoras.



EXP-UNC: 60737/2017

### 1.5. Aprendizaje no supervisado

Aprendizaje no supervisado y semi supervisado. Clustering. Kernel PCA. Matrix Factorization. Modelos con variables latentes. LDA. LSI.

Arquitecturas neuronales para extracción de features. Autoencoders. Embeddings.

Sistemas de recomendación. Collaborative filtering.

## 2. Materias Optativas

- 2.1. Infraestructura para grandes volúmenes de datos
- 2.2. Programación distribuida sobre grandes volúmenes de datos
- 2.3. Análisis de grandes volúmenes de datos
- 2.4. Aprendizaje con redes neuronales
- 2.5. Text Mining
- 2.6. Conceptos de visión por computadoras
- 2.7. Bioinformática
- 2.8. Astroestadística
- 2.9. Modelos Gráficos Probabilísticos
- 2.10. Aprendizaje por Refuerzo
- 2.11. Aplicaciones en Economía
- 2.12. Análisis GeoEspacial
- 2.13. Internet de las Cosas
- 2.14. Aplicaciones en Robótica

## COMPETENCIAS A DESARROLLAR

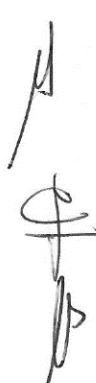
Manejo concreto en casos prácticos de herramientas de análisis estadístico y visualización (por ejemplo, R, Octave, Google Analytics).

Manejo concreto en casos prácticos de frameworks de aprendizaje automático supervisado y no supervisado (por ejemplo, weka, scikit-learn, keras, tensorflow, coffee).

Comunicar de manera efectiva un análisis de problema, sus principales dificultades, riesgos y complejidad de la solución propuesta. Análisis comparativo de soluciones alternativas.

Realizar estimaciones de costo de soluciones para problemas basados en datos.

Conocer en profundidad los diferentes métodos y métricas de evaluación, aplicar las más adecuadas en cada caso concreto e interpretar los resultados. Comunicar de forma efectiva esta evaluación.






Universidad  
Nacional  
de Córdoba



Facultad de Matemática,  
Astronomía, Física y  
Computación

EXP-UNC: 60737/2017

## METODOLOGÍA



La diplomatura tendrá un enfoque teórico-práctico. Se presentarán los conceptos teóricos con una fuerte motivación en problemas y soluciones que se encuentran en el entorno, y se realizarán ejercicios y pequeños proyectos en los que los estudiantes pondrán en práctica los conceptos teóricos. Estos ejercicios se realizarán como parte de las horas presenciales, complementando las clases magistrales, y también en las horas virtuales, con asistencia por parte de tutores a distancia.





Universidad Nacional de Córdoba



Facultad de Matemática, Astronomía, Física y Computación

EXP-UNC: 60737/2017

**GRILLA PARA PREVER LA DISTRIBUCIÓN DE METODOLOGÍAS DOCENTE EN HORAS Y LUEGO "TRADUCIRLAS " A RTF**

	Horas dirigidas por el Profesor		Horas de Trabajo Autónomo	HORAS TOTALES
	Horas presenciales	Horas no presenciales		
Actividades Introdutorias				
Sesiones magistrales	60			60
Eventos científicos				
Prácticas	En el aula			
	En el laboratorio			
	En aula de informática	60		60
	De campo			
	De observación			
Practicum				
Prácticas externas				
Seminarios				
Exposiciones				
Debates				
Tutorías				
Actividades de seguimiento on line		10		10
Preparación de trabajos				
Otros trabajos en la casa				
Resolución de problemas			50	50
Estudio de casos				
Foros de discusión				
Pruebas orales				
Pruebas de desarrollo				
Pruebas en situación auténtica				
<b>TOTAL</b>	<b>120</b>	<b>10</b>	<b>50</b>	<b>180</b>
	<b>RTF</b>			<b>3</b>

*Handwritten signature or initials.*



EXP-UNC: 60737/2017

### DISTRIBUCIÓN DEL TOTAL DE RTF

Número de horas reloj dirigidas por el profesor	130
Número de horas reloj de trabajo AUTÓNOMO (estudio individual y grupal, preparación y elaboración del trabajo final/ evaluaciones parciales y finales)	50
Total horas reloj	180
Total RTF	3

### EVALUACIÓN

Las materias se dictarán en forma consecutiva, y cada una de ellas tendrá una evaluación formativa en un momento intermedio de la materia y de tipo diagnóstico al final de cada materia. Se tomarán mediante una prueba estructurada escrita y una defensa oral opcional, a criterio del profesor. Se aprobará con 60%.

### RECONOCIMIENTO DE LOS TRAYECTOS FORMATIVOS (RTF)

No se establecen reconocimientos.

### BIBLIOGRAFÍA

Pedro Domingos. The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World, Basic Books, 2015.

Ian Goodfellow, Yoshua Bengio and Aaron Courville, Deep Learning, MIT Press, 2016.

Trevor Hastie, Robert Tibshirani, Jerome H. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, (Second Edition), Springer Verlag, 2009.  
Tom Mitchell, Machine Learning, McGraw Hill, 1997.

Ian Witten, Eibe Frank and Mark Hall, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 2011.





Universidad  
Nacional  
de Córdoba



Facultad de Matemática,  
Astronomía, Física y  
Computación

EXP-UNC: 60737/2017

**CRONOGRAMA**

Contenido	Fecha	Semana/sesión
Análisis y visualización de datos	13, 14, 27, 28/04/2018	1 y 2
Análisis exploratorio y curación de datos	4, 5, 18, 19/05/2018	3 y 4
Introducción al aprendizaje automático	1, 2, 15, 16/06/2018	5 y 6
Aprendizaje supervisado	29, 30/06, 27, 28/07/2018	7 y 8
Aprendizaje no supervisado	10, 11, 24, 25/08/2018	9 y 10
Bloque de optativas 1	7, 8, 28, 29/09/2018	11 y 12
Bloque de optativas 2	12, 13, 26, 27/10/2018	13 y 14
Bloque de optativas 3	9, 10, 23, 24/11/2018	15 y 16

A  
F  
L



Universidad  
Nacional  
de Córdoba



Facultad de Matemática,  
Astronomía, Física y  
Computación

EXP-UNC: 60737/2017

### MODELO DE CERTIFICADO

La Facultad de MATEMÁTICA, ASTRONOMÍA, FÍSICA Y COMPUTACIÓN de la UNIVERSIDAD NACIONAL DE CÓRDOBA **certifica** que ....., DNI N°..... ha finalizado la DIPLOMATURA EN CIENCIA DE DATOS, APRENDIZAJE AUTOMÁTICO Y SUS APLICACIONES, aprobada por RESOLUCIÓN CD N° ....., con una carga horaria de 180 horas.

Firma

Firma

Córdoba, .. de ..... de 2018

**El presente certificado no es habilitante para el ejercicio profesional.**