

ANEXO

Diplomatura en Ciencia de Datos, Aprendizaje Automático y sus Aplicaciones -- Cohorte 2023 --

1. Datos generales

- a. Tipo de curso: Diplomatura.
- b. Certificado: Diplomatura en Ciencia de Datos, Aprendizaje Automático sus Aplicaciones.
- c. Unidad Académica: FAMAF.
- d. Período lectivo: abril a diciembre de 2023.
- e. Disponibilidad horaria del alumnado: 195 horas totales. 120 horas virtuales sincrónicas y aproximadamente 75 horas de trabajo práctico asincrónico. Las 120 horas virtuales sincrónicas corresponden a clases teórico-prácticas de 5 materias obligatorias (16 horas cátedra cada una), dos materias optativas de 16 horas cada una y una materia optativa de 8 horas. El tiempo estimado necesario para realizar los trabajos prácticos para aprobar la Diplomatura es de 8 horas por cada materia de 16 horas, 4 horas por la materia optativa de 8 horas y 15 horas de consultas por mail de mentoría.
- f. Horario de clases sincrónicas virtuales: viernes de 18 a 22 horas y sábados de 10 a 14 horas.
- g. Lugar en que se desarrollarán las clases: Aula Virtual - FAMAF.
- h. Número mínimo y máximo de estudiantes: Mínimo 112, máximo 160.
- i. Perfil del alumnado que pueden asistir al curso: El curso orientado principalmente a estudiantes y profesionales de la computación, ingenierías o de actividades que involucren el trabajo y modelado estadístico de datos, así como a profesionales de otras áreas con un buen conocimiento funcional de programación.

2. Programa

2.1. Fundamentación

En los últimos años, el área de ciencia de datos ha desarrollado ramificaciones de gran importancia, como el aprendizaje automático, la inteligencia artificial y múltiples aplicaciones industriales. Este desarrollo ha generado una gran demanda en todos los ámbitos socio-productivos y académicos, y a la vez ha impulsado avances cada vez más importantes en las mismas áreas. Esta demanda sólo 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 actualización permanente. Al mismo tiempo, resulta ineludible basarse en una rigurosa metodología empírica, basada en experimentación y evaluación. Es por esta razón que se propone una diplomatura que proveerá el vínculo entre fundamentos teóricos y prácticos para el análisis de problemas y diseño de soluciones basadas en el procesamiento de datos.

Los y las 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.

2.2. Objetivos

- Proveer a los y las 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.
- Proveer formación a egresados/as de la diplomatura y a estudiantes avanzados/as en el desarrollo de capacidades de dirección de proyectos en ciencia de datos, a través de un programa de mentorías a estudiantes en su estadio inicial.

2.3. Contenidos

2.3.1. Materias Obligatorias

Cada materia obligatoria consta de 16 horas de clases teórico-prácticas virtuales sincrónicas y un estimativo de 8 horas necesarias para realizar los trabajos prácticos en grupos interdisciplinarios de hasta 5 integrantes, supervisadas por docentes de seguimiento de manera virtual.

Análisis y visualización de datos

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

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. Reducción de dimensionalidad. PCA. Sesgo en diseño, sistematización de información y muestreo. Sesgo estadístico. Detección de anomalías. Detección de outliers. Detección de ruido, errores, valores faltantes. Normalización.

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

Aprendizaje supervisado

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

Aprendizaje no supervisado

Aprendizaje no supervisado y semi supervisado. Clustering clásico, K-medias, Mean Shift, Jerárquicos. Proyecciones. Kernel PCA. Matrix Factorization. Modelos con variables latentes. LDA. LSI. Arquitecturas neuronales para extracción de features. Autoencoders. Embeddings. Sistemas de recomendación. Collaborative filtering.

2.3.2. Mentorías

Como parte de los trabajos prácticos, los y las estudiantes llevarán a cabo un proyecto integrador transversal a todas las materias obligatorias. El proyecto se realizará en grupos de no más de 4 personas. Este proyecto transversal busca presentar al alumnado un problema real que deberá abordar a lo largo del cursado, utilizando las herramientas brindadas por cada materia obligatoria. El grupo estará a cargo de un mentor especialista en el problema, quien le asignará una lista de tareas diseñada para abordar cada problema de manera paulatina a medida que se cursa la diplomatura. La lista será diseñada con el apoyo de los/as coordinadores/as de mentorías. Se busca que el grupo desarrolle habilidades para abordar problemas concretos y que interactúe con el/la mentor/a especialista. El/la mentor/a no solo deberá poseer conocimiento específico del problema a resolver, sino también de las herramientas brindadas en las materias optativas. En este contexto el/la mentor/a se encargará de describir los objetivos y las tareas a realizar por parte del grupo, y asistirlo en los problemas que surjan. Cada mentoría supone 3 horas de interacción grupo - mentor por materia obligatoria.

2.3.3. Materias Optativas

Las materias optativas se determinarán en el mes de julio de 2023, de la siguiente manera:

- Mayo: Llamado a presentación de propuestas de materias optativas.
- Julio: Selección las materias optativas a dictarse, según disponibilidad presupuestaria y a criterio del Consejo Académico Profesional de la Diplomatura.
- Agosto: Inscripción, a las materias optativas, de los/as estudiantes inscriptos en la Diplomatura.
- Agosto - Noviembre: Inscripción de otros/as estudiantes a las materias optativas (aranceladas).

Cada estudiante deberá cubrir 40 horas teórico-prácticas sincrónicas de materias optativas y un estimativo de 20 horas para la realización de los trabajos prácticos de materias optativas, a elección entre la oferta del año en que curse.

3. Estructura de dictado

El dictado de la Diplomatura en Ciencias de Datos, Aprendizaje Automático y sus Aplicaciones 2023 se realizará en formato virtual, de la siguiente forma

- Clases teórico-prácticas virtuales sincrónicas de 16 horas en 4 encuentros de 4 horas cada uno, los días viernes de 18 a 22 hs. y sábados de 10 a 14 hs., semana de por medio, con los/as docentes presentes interactuando con los/as estudiantes, dictadas usando la plataforma Google Meet siguiendo el cronograma detallado a continuación.
- Cada materia tendrá al menos un trabajo práctico asociado para resolver en forma sincrónica usando las herramientas discutidas en clase. La resolución de estos trabajos

se realizará obligatoriamente en equipos de hasta 5 integrantes formados por la coordinación de la Diplomatura. En cada grupo se priorizará la diversidad de formaciones favoreciendo trabajo interdisciplinario.

- Las devoluciones de los trabajos prácticos asincrónicos, corregidos por los/as docentes de seguimiento, serán realizadas mediante la plataforma Google Meet en los fines de semana alternos sin clases sincrónicas teórico-prácticas, en horarios a convenir con los/as estudiantes.
- La comunicación general de los/as docentes con los/as estudiantes se realizará mediante la plataforma Slack y la comunicación administrativa a través del email de la diplomatura.
- Las devoluciones de los/as mentores/as sobre los trabajos preparados sobre los grupos de datos de mentoría serán realizadas también mediante la plataforma Google Meet.
- Las presentaciones finales de las mentorías se realizarán mediante videos asincrónicos, mientras que la defensa de dichas mentorías será realizada en forma sincrónica mediante la plataforma Google Meet.
- La coordinadora General y la coordinadora de mentorías estarán a cargo de recolectar las notas de aprobación de todos los trabajos prácticos y mentorías en un único documento de seguimiento de cada estudiante, el cual también tendrá asociado un seguimiento del pago de las cuotas que correspondan.

4. Evaluación

La evaluación se realizará mediante trabajos prácticos en grupo, desarrollados tanto durante la cursada como también de forma asincrónica. Dichos trabajos prácticos se entregarán y serán evaluados por los/as docentes responsables de los trabajos prácticos.

Para obtener el Diploma de la Diplomatura en Ciencia de Datos, Aprendizaje Automático y sus Aplicaciones, las condiciones mínimas establecidas son:

1. Asistir al 80% de las clases.
2. Aprobar 2 materias optativas de 16 horas y una de 8 horas (medio módulo), con criterios de aprobación específicos para cada materia.
3. Aprobar el 60% de los trabajos prácticos (8 trabajos prácticos), con las siguientes condiciones de aprobación:
 - a. Cubrir los ítems propuestos por la consigna,
 - b. Entregar el trabajo en la fecha propuesta por el/la docente o mentor/a,
 - c. Incorporar las observaciones y correcciones, si las hubiera,
 - d. Entregar el trabajo con las correcciones en la fecha propuesta por el profesor o mentor.

De los 8 trabajos prácticos a aprobar, por lo menos 3 deberán corresponder a mentorías

- y los otros cinco a materias obligatorias diferentes.
4. Presentar el trabajo de síntesis de las mentorías:
 - a. Un informe que se publicará en la página de la Diplomatura, cubriendo los ítems y observaciones propuestos por el/la mentor/a, y reflejando los resultados de algunos de los trabajos prácticos relacionados, a presentar en el mes de octubre.
 - b. Una presentación oral en el mes de noviembre.

5. Cronograma Propuesto

Actividad	Fecha
Período de pre-inscripción	28/2 - 14/3
Elaboración del orden de méritos	15/3 - 21/3
Publicación del orden de méritos	22/3
Inscripción	23/3–31/3
Análisis y visualización de datos	14/4, 15/4, 21/4, 22/4
Análisis exploratorio y curación de datos	12/5, 13/5, 19/5, 20/5
Introducción al aprendizaje automático	2/6, 3/6, 9/6, 10/6
Aprendizaje supervisado	23/6, 24/6, 7/7, 8/7
Receso Invernal	10/7–22/7
Aprendizaje no supervisado	28/7, 29/7, 4/8, 5/8
Colación de grados	15 de diciembre

6. Cuerpo Académico Propuesto

La selección del cuerpo académico se realizará según el siguiente proceso.

- Para las materias obligatorias, se convocará a docentes del año anterior con evaluaciones positivas a cubrir el mismo cargo del año anterior. Si quedaran vacantes, se realizará un llamado abierto para cubrirlos. Con los/as postulantes del llamado, el/la coordinador/a realizará una propuesta al Consejo Académico Profesional, el cual elevará la propuesta a la Secretaría Académica de FAMAF.
- Se planea realizar este llamado el día 22 de febrero 2023, el tribunal estará compuesto por la coordinadora, la Dra. Laura Alonso Alemany y el Dr. Oscar Reula.
- Para los proyectos de mentoría, se realizará un llamado abierto en el mes de marzo. Con los/as postulantes y proyectos del llamado, el coordinador hará una propuesta al Consejo Académico Profesional. Se ofertarán los proyectos a los/as estudiantes, y aquellos que cubran con un mínimo de 2 estudiantes seleccionados podrán realizarlo.
- Para las materias optativas, se realizará un llamado abierto en el mes de mayo. Con los/as postulantes del llamado, y conforme a disponibilidad presupuestaria, el/la coordinador/a hará una propuesta al Consejo Académico Profesional, que elevará la propuesta de materias optativas a la Secretaría Académica de FAMAF.

6.1. Materias Obligatorias

El objetivo pedagógico primordial de la Diplomatura es la construcción colectiva e interdisciplinar de las soluciones, lo cual se desarrolla desde el mismo dictado con dos o más docentes interactuando en el aula en todo momento, en la construcción del trabajo entregable de cada materia, que se resuelve en grupo y en la mentoría anual, también en grupo.

Es tarea de los/as docentes de las materias de teórico-prácticas obligatorias

- preparar 16 horas de material teórico-práctico para ser dictado en forma virtual sincrónica,
- preparar un material práctico que deberá ser realizado por los/as estudiantes y entregado a los/as docentes de práctico para su corrección y la correspondiente aprobación de la materia,
- poner a disposición de los/as estudiantes el material didáctico de la materia preparado por los/as docentes en el repositorio libre Github, la página web de la Diplomatura y el Calendario anual, junto a la grabación de la clase,
- interactuar con los/as estudiantes en la plataforma Slack, en el canal de la materia habilitado para tal fin,
- prestar asistencia a los/as docentes de práctico para alcanzar un criterio común de corrección del material entregado por los/as estudiantes para la aprobación de la materia,

- estar presentes en las 16 horas de dictado, colaborando activamente con los/as estudiantes y con capacidad de retomar la clase si el/la docente que está dictando a un determinado momento tiene problemas con la conexión virtual.

Cada docente percibirá por las tareas descritas anteriormente honorarios por 16 horas de trabajo.

6.1.1. Coordinación de Prácticos

El Coordinador de trabajos prácticos armará el listado de notas de los/las estudiantes de todas las materias de la Diplomatura, y apoyará a los/as docentes de práctico en el armado de los grupos de trabajo. El/la coordinador/a de prácticos tendrá una carga horaria de 55 horas totales y cobrará en cuatro cuotas, en los meses de mayo, julio, septiembre y diciembre.

6.2. Materias optativas

Las materias optativas para el año 2023 se determinarán en el mes de julio de 2023, según disponibilidad presupuestaria, a partir de la convocatoria abierta para propuesta de materias optativas que se realizará en el mes de mayo, y de la selección de los/as estudiantes incluida en el formulario de inscripción.

6.3. Mentorías

Habrá un cuerpo de mentores/as que realizará tareas de seguimiento de proyectos de los/as estudiantes. Habrá una convocatoria abierta para mentores/as, de los cuales el Consejo Académico Profesional seleccionará a todos/as aquellos/as que cumplan con los requisitos básicos para ser mentores/as: formación comprobable en ciencia de datos y propuesta de proyecto según el pedido de la convocatoria, priorizando aquellos/as que no hayan sido mentores/as en la Diplomatura en años anteriores.

Estos/as mentores/as serán propuestos en el mes de marzo/abril para su nombramiento por parte del Consejo Directivo de FAMAF. Después, se ofrecerán los diferentes proyectos a los/as estudiantes de la Diplomatura y se efectivizarán aquellos proyectos que sean seleccionados por un mínimo de 2 estudiantes.

El proyecto de mentoría ocupará 20 horas de trabajos prácticos no presenciales. Los/as mentores/as tendrán una carga horaria de 20 horas por un grupo y 35 horas por dos grupos.

6.3.1. Coordinación de Mentorías

En el año 2020 se introdujo la coordinación de mentorías, figura que supuso una importante mejora en el programa de mentorías. El/la coordinador/a de mentorías tendrá una carga horaria de 10 horas por mes durante 10 meses (de febrero a noviembre) y cobrará en cuatro cuotas, en los meses de mayo, julio, septiembre y diciembre.

6.4. Coordinación general

El/la Coordinador/a General tendrá una carga horaria de 10 horas por mes durante 10 meses (de febrero a noviembre) y cobrará en cuatro cuotas, los meses de mayo, julio, septiembre y diciembre.

7. Selección de estudiantes

Requisitos de ingreso:

- Conocimientos de programación a nivel instrumental, especialmente comprensión y programación en Python y uso de Jupyter notebooks.
- Conocimientos de probabilidad y estadística al nivel de un curso de grado.
- Conocimientos de álgebra al nivel de un curso de grado.

Se proveerán materiales de autodiagnóstico y nivelación para que los/as postulantes sin formación curricular, o sin práctica reciente, puedan determinar si cumplen con los requisitos de ingreso para la Diplomatura y puedan administrar los contenidos necesarios para complementar su formación de base, si fuera necesario.

En el caso de que el número de inscripciones supere el número de vacantes, el Consejo Académico Profesional de la Diplomatura realizará un orden de méritos que publicará oportunamente.

8. Aranceles

El arancel es de \$400.000 (pesos cuatrocientos mil) total, el cual podrá abonarse en hasta 5 cuotas mensuales, iguales y consecutivas de \$80.000 (pesos ochenta mil) una vez aceptada la inscripción de los/as estudiantes, y no más allá del día 10 de agosto de 2023. Las demoras en el pago tendrán una penalización del 5% por cada 10 días de demora. El pago de la primera cuota supondrá la efectivización de la inscripción en la Diplomatura.

Podrán acceder a inscripción reducida de \$300.000 (pesos trescientos mil) estudiantes no universitarios, trabajadores de empresas micro y pequeñas y trabajadores de empresas del Córdoba Technology Cluster. Se ha previsto un cupo de medias becas (con un arancel de \$200.000) y becas completas, para trabajadores del sector público, estudiantes regulares de universidades públicas, trabajadores de la UNC y casos excepcionales.

9. Honorarios

Los honorarios para las materias obligatorias de la presente cohorte:

- Clases teóricas: \$10.000 por hora.
- Prácticas asistidas, correcciones y devoluciones de trabajos prácticos: \$8.000 por hora.
- Mentoría: \$6.000 por hora.

Los honorarios de las materias optativas podrán ser actualizados por inflación de acuerdo con las posibilidades de presupuesto y cantidad de estudiantes que cursen.

Los honorarios de coordinación general y de mentoría serán de \$7.500 por hora y los de coordinación de prácticos serán de \$5.000 por hora.

Se prevé incorporar la figura de un/a facilitador/a responsable de apoyo administrativo a la inscripción general y de materias obligatorias, manejo de los canales de comunicación, formularios Google, con una carga horaria total de 15 horas semanales y honorarios de \$1.600 por hora, durante 10 meses.