

PLAN DE ESTUDIO LICENCIATURA EN CIENCIA DE DATOS

I- IDENTIFICACIÓN DE LA CARRERA

a) Nombre de la carrera: Licenciatura en Ciencia de Datos

b) Fundamentación:

La Universidad Nacional del Delta concibe la educación como pilar fundamental para la construcción de la Argentina del futuro donde la ciencia, la tecnología y la innovación se encuentren articuladas entre el nivel secundario con el superior para lograr progresivamente una mejor formación de quienes desean acceder a la universidad. Tiene como objetivo contribuir al desarrollo del conocimiento, promover el talento humano, el empleo de calidad y la producción vinculando el sistema productivo y social. Surge así, una nueva propuesta educacional de excelencia que orienta su misión institucional y de formación en virtud del proyecto que le dio origen y que pretende promover la democratización del conocimiento, los valores humanísticos, a través de un modelo educativo que facilite la formación de profesionales competentes para transformar la realidad con responsabilidad social y con un fuerte sentido de solidaridad.

La Universidad Nacional del Delta tiene por funciones:

- Vincular entre el ámbito académico y las organizaciones territoriales para la formulación de políticas públicas inclusivas, participativas y pensadas desde un enfoque de derechos.
- Contribuir al desarrollo socioeconómico territorial generando conocimiento mediante la investigación, acerca de la realidad política local, nacional y regional para formular respuestas a los problemas sociales, sanitarios y comunitarios.
- Contribuir al desarrollo equitativo de la Argentina mediante la implementación de estrategias educativas virtuales y el trabajo conjunto con el resto de las universidades nacionales.

Esta casa de estudios busca preparar jóvenes para los trabajos del siglo XXI contribuyendo al acceso e inclusión, acrecentando su nivel educativo y desarrollo personal y profesional, para beneficio propio y de la sociedad, dando respuesta a las necesidades de la comunidad de cada uno de los municipios y su zona de influencia.

En las últimas dos décadas, nuevas innovaciones en información y tecnología han revolucionado la gestión, el estudio social y la política contemporáneas. Los avances en la capacidad computacional han sido acompañados por una expansión nunca antes vista en la cantidad de información disponible. Cada día se producen y diseminan cientos de miles de datos creando demandas de almacenamiento, procesamiento, minado e interpretación nunca antes vistas.

Grandes compañías están minando datos para capitalizarse, estudios de abogados están minando datos para diseñar sus estrategias legales y muchos Estados han incorporado políticas de distribución de datos abiertos que requieren la administración de bases masivas de datos. Para poder afrontar estos grandes volúmenes de información es indispensable formar recursos humanos capacitados para poder administrar y gestionar toda esta información y, simultáneamente, poder diseñar políticas que tenga impacto real en el territorio.

La integración de grandes volúmenes de datos en los procesos de gobierno y administración en todas las escalas geográficas, han transformado profundamente las herramientas de análisis, los perfiles de capacitación necesarios, los estándares de investigación, la gestión pública y la territorial.

En efecto, la existencia y disponibilidad de grandes bases de datos asociadas al desarrollo de la infraestructura de hardware y software necesarias para su procesamiento y análisis; ha posibilitado la difusión de modelos de intervención y gestión inteligente del territorio basados en la minería de datos y la predicción de problemáticas, necesidades y comportamientos geográficos por medio del uso intensivo de técnicas de aprendizaje automatizado.

En la actualidad la cantidad de datos disponibles crecen más rápido que la capacidad de analizarlos. En este marco, se necesitan profesionales con la capacidad de explotarlos y descubrir información (conocimiento) de manera eficiente. El aumento sostenido que se espera en la demanda nacional y global de servicios asociados a la explotación de datos y descubrimiento del conocimiento (Ciencia de Datos), necesita dotarse de recursos humanos altamente especializados que cubran los aspectos integrales del sector productivo en estas áreas.

La Argentina cuenta con varios de los factores necesarios para aprovechar este potencial. Particularmente en el ámbito de desarrollo de modelos estadísticos, entre ellos una amplia base de empresas del sector de distintas características y tamaños que trabajan tanto en el mercado local como en el internacional y una cantidad interesante de profesionales con capacidades competitivas a nivel global.

Es en este contexto que la Universidad Nacional de Delta se propone aportar al sector público y privado, recursos humanos en los que se destaque la capacidad de proveer servicios de alta calidad. Esta visión motiva la orientación del plan, aspirando a conjugar una práctica extensiva en habilidades directamente relacionadas con las necesidades que percibimos en el mercado laboral, con una sólida formación en los conceptos de Ciencia de Datos.

II- OBJETIVOS DE LA CARRERA

a) Objetivos Generales:

- Formar profesionales y recursos humanos capaces de concebir soluciones a un amplio espectro de problemas asociados a las tareas de diseño, e implantación de modelos de

Ciencia de Datos; capaces de explotar datos y descubrir conocimiento en el estado del arte de la tecnología, con pensamiento respetando los factores legales, éticos, ambientales y de seguridad, con un fuerte y sólido conocimiento de la realidad nacional y social.

b) Objetivos Específicos

1. Formar profesionales en el campo de la informática con un fundamento sólido en la implementación de modelos de Ciencia de Datos tanto de aprendizaje supervisado como no supervisado.
2. Formar recursos humanos con pensamiento crítico y conciencia social, con capacidades para vincularse con la industria del software regional, nacional e internacional.
3. Generar una propuesta de enseñanza que permita alcanzar capacidades de participar en proyectos vinculados a Ciencia de Datos.
4. Proporcionar un marco formativo integral referido al objeto de estudio de la estadística, la informática, la matemática y la ciencia de datos.
5. Favorecer la orientación hacia el asesoramiento en los diferentes campos de estudio relacionados

III- CARACTERÍSTICAS DE LA CARRERA

- a) Nivel académico de la carrera: Grado
- b) Especificación de la modalidad: Presencial
- c) Duración de la carrera: 8 cuatrimestres
- d) Asignación horaria total de la carrera (en horas reloj): 2.144 horas
- e) Nombre del/os Título/s a otorgar: Licenciado/a en Ciencia de Datos

IV- PERFIL DEL GRADUADO/A

El/la Licenciado/a en Ciencia de Datos de la Universidad Nacional de Delta está enfocado en dar respuestas a necesidades de la sociedad, empresas y organismos a través de procesos de puesta en funcionamiento de modelos de Ciencia de Datos ya sea desde la construcción de los mismos, como así también desde la adaptación de soluciones existentes, principalmente del ámbito del software libre. Posee una gran formación ética profesional, y una estrecha relación con el sector productivo regional. Con alta capacidad de desarrollo de proyectos propios, está preparado para intervenir en el ámbito público y privado.

El/la Licenciado/a en Ciencia de Datos es un profesional universitario que tendrá saberes capacidades y habilidades para:

- Manipular, explorar, limpiar y preparar cualquier fuente de datos antes de su procesamiento y realizar análisis exploratorios del conjunto de datos relevantes para un problema.



- Identificar patrones y tendencias de los datos que servirán para informar modelos estadísticos.
- Desarrollar modelos estadísticos para el análisis y la resolución de problemas inéditos en su ámbito de desarrollo profesional.
- Utilizar, programar y parametrizar herramientas Informáticas de estadística avanzada para el análisis de datos.
- Evaluar críticamente las soluciones propuestas o implementadas en el área de la Ciencia de Datos e identificar posibles formas de mejorarlas.

V- ALCANCES DEL TÍTULO

Los alcances del Licenciado/a en Ciencia de Datos son:

- Gestionar proyectos de relevamiento de problemas del mundo real, a partir de aplicaciones referentes a las Ciencias de Datos.
- Dirigir equipos de trabajo para el relevamiento y análisis de los procesos funcionales de una Organización, con la finalidad de establecer proyectos de diseño aplicados a la Ciencia de Datos.
- Asistir en los estudios técnico-económicos de factibilidad y/o referentes a la configuración y dimensionamiento del almacenamiento y uso de datos.
- Intervenir en proyectos tanto de Aprendizaje Supervisado como de Aprendizaje No Supervisado.
- Participar en tareas de auditoría de los sistemas de procesamiento de datos
- Participar en tareas de investigación científica básica y aplicada en temas de Ciencias de Datos.
- Formar parte de equipos multidisciplinarios que integren soluciones de informática.
- Desarrollar actividades de docencia e investigación en instituciones educativas de altos estudios.

VI- CONDICIONES DE INGRESO

Acreditar estudios secundarios completos y haber cursado satisfactoriamente el Curso de Preparación Universitaria (CPU). Excepcionalmente, los mayores de 25 años que no posean título secundario, según lo establece el Artículo 7 de la Ley de Educación Superior 24.521, podrán ingresar siempre que demuestren los conocimientos necesarios a través de las evaluaciones que realice la Universidad.

VII- DISEÑO Y ORGANIZACIÓN CURRICULAR

La carrera está conformada por 4 campos de formación que se complementan y articulan:

1. Campo de Formación Común (CFC).
2. Campo de Formación Básica (CFB).

3. Campo de Formación Disciplinar (CFD).
4. Campo de la Práctica Profesional (CPP).

1. CAMPO DE FORMACIÓN COMÚN

Todas las carreras de la Universidad Nacional del Delta comparten un Campo de Formación Común (CFC). Esto se constituye a partir de un conjunto de materias obligatorias que se dictan en todas las carreras de esta alta casa de estudios. El campo de formación común buscará la inserción de los estudiantes a la vida académica universitaria con una base necesaria para el comienzo de sus actividades en la UNDelta.

El CFC comprende las siguientes asignaturas:

- 1.1 Alfabetización Digital Académica
- 1.2 Cultura General en la Universidad
- 1.3 Desafíos y Problemáticas del Delta.
- 1.4 Inglés I
- 1.5 Inglés II

2. CAMPO DE FORMACIÓN BÁSICA

El Campo de Formación Básica (CFB) es un espacio curricular que propone un estudio gradual de contenidos para el comienzo del perfil del egresado/a, donde definirá y desarrollará las competencias y habilidades fundamentales para el éxito académico y profesional.

El CFB comprende las siguientes asignaturas:

- 2.1 Introducción a la Ciencia de Datos
- 2.2 Introducción al Aprendizaje Automático
- 2.3 Análisis I
- 2.4 Programación I
- 2.5 Análisis II
- 2.6 Probabilidad y Estadística
- 2.7 Estadística e Inferencia I
- 2.8 Programación II
- 2.9 Estadística e Inferencia II
- 2.10 Inglés III
- 2.11 Inglés IV

3. CAMPO DE LA FORMACIÓN DISCIPLINAR

Este campo incluye saberes propios y específicos de la formación. Desarrolla conocimientos prácticos y teóricos acerca de las diferentes áreas de trabajo, así como también saberes y destrezas orientados a la educación, investigación y gestión.

El CFD comprende las siguientes asignaturas:

- 3.1 Algoritmos y Estructuras de Datos I
- 3.2 Ciencia de Datos
- 3.3 Infraestructura para Ciencia de Datos
- 3.4 Base de Datos
- 3.5 Algoritmos y Estructuras de Datos II
- 3.6 Aprendizaje Automático
- 3.7 Electiva I
- 3.8 Aprendizaje Profundo
- 3.9 Electiva II

4. CAMPO DE LA PRÁCTICA PROFESIONAL

Se espera que a través de este espacio los estudiantes logren relacionarse con el campo de conocimiento, de inserción laboral y de desarrollo profesional, desde el inicio de sus trayectorias universitarias, mediante la realización de diferentes experiencias de acercamiento, inserción y participación en instituciones y/u organizaciones del territorio.

El CPP comprende las siguientes asignaturas:

- 4.1 Práctica Profesional I
- 4.2 Práctica Profesional II
- 4.3 Taller de Análisis e Investigación de la Práctica Profesional

VIII- ESTRUCTURA CURRICULAR: CAMPO FORMATIVO, ASIGNATURAS, HORAS SEMANALES Y TOTALES, RÉGIMEN DE CURSADO, CORRELATIVIDADES Y MODALIDAD

Campo Formativo	Asignatura	hs semanales	hs totales	Régimen de Cursada	Correlatividades *
1 - Formación Común	1.1 Alfabetización Digital Académica	2	32	Cuatrimestral	
	1.2 Cultura General en la Universidad	2	32	Cuatrimestral	
	1.3 Desafíos y Problemáticas del Delta	2	32	Cuatrimestral	
	1.4 Inglés I	2	32	Cuatrimestral	
	1.5 Inglés II	2	32	Cuatrimestral	1.4
2 - Formación Básica	2.1 Introducción a la Ciencia de Datos	6	96	Cuatrimestral	
	2.2 Introducción al Aprendizaje Automático	6	96	Cuatrimestral	
	2.3 Análisis I	6	96	Cuatrimestral	
	2.4 Programación I	6	96	Cuatrimestral	2.1
	2.5 Análisis II	4	64	Cuatrimestral	2.3
	2.6 Probabilidad y Estadística	6	96	Cuatrimestral	2.1
	2.7 Estadística e Inferencia I	4	64	Cuatrimestral	2.6
	2.8 Programación II	6	96	Cuatrimestral	2.4
	2.9 Estadística e Inferencia II	6	96	Cuatrimestral	2.7
	2.10 Inglés III	3	48	Cuatrimestral	1.5
	2.11 Inglés IV	3	48	Cuatrimestral	2.9
3 - Formación Disciplinar	3.1 Algoritmos y Estructuras de Datos I	6	96	Cuatrimestral	
	3.2 Ciencia de Datos	6	96	Cuatrimestral	2.1
	3.3 Infraestructura para Ciencia de Datos	4	64	Cuatrimestral	2.1
	3.4 Base de Datos	6	96	Cuatrimestral	3.1
	3.5 Algoritmos y Estructuras de Datos II	6	96	Cuatrimestral	3.1
	3.6 Aprendizaje Automático	6	96	Cuatrimestral	3.1
	3.7 Electiva I	6	96	Cuatrimestral	
	3.8 Aprendizaje Profundo	6	96	Cuatrimestral	3.6
	3.9 Electiva II	6	96	Cuatrimestral	
4 - Práctica Profesional	4.1 Práctica Profesional I	4	64	Cuatrimestral	2.2, 2.3 y 2.4
	4.2 Práctica Profesional II	6	96	Cuatrimestral	4.1
	4.3 Taller de Análisis e Investigación de la Práctica Profesional	6	96	Cuatrimestral	4.2

* Correlativa cursada en condición regular para CURSAR. Correlativa APROBADA para promocionar y/o rendir examen final



IX- RÉGIMEN DE CURSADO POR AÑO Y CUATRIMESTRE

Primer Año: Primer Cuatrimestre		
Asignatura	hs semanales	hs totales
1.1 Alfabetización Digital Académica	2	32
2.1 Introducción a la Ciencia de Datos	6	96
2.3 Análisis I	6	96
1.4 Inglés I	2	32
Horas total primer cuatrimestre		256
Primer Año: Segundo Cuatrimestre		
Asignatura	hs semanales	hs totales
3.1 Algoritmos y Estructura de Datos I	6	96
2.2 Introducción al Aprendizaje Automático	6	96
1.2 Cultura General en la Universidad	2	32
1.5 Inglés II	2	32
Horas total segundo cuatrimestre		256
Segundo Año: Tercer Cuatrimestre		
Asignatura	hs semanales	hs totales
2.5 Análisis II	4	64
2.4 Programación I	6	96
1.3 Desafíos y Problemáticas del Delta	2	32
2.9 Inglés III	3	48
Horas total tercer cuatrimestre		240
Segundo Año: Cuarto Cuatrimestre		
Asignatura	hs semanales	hs totales
3.2 Ciencia de Datos	6	96
2.6 Probabilidad y Estadística	6	96
2.10 Inglés IV	3	48
Horas total cuarto cuatrimestre		240
Tercer Año: Quinto Cuatrimestre		
Asignatura	hs semanales	hs totales
2.7 Estadística e Inferencia I	4	64
2.8 Programación II	6	96
3.3 Infraestructura para Ciencia de Datos	4	64
4.1 Práctica Profesional I	4	64
Horas total quinto cuatrimestre		288



Tercer Año: Sexto Cuatrimestre		
Asignatura	hs semanales	hs totales
2.9 Estadística e Inferencia II	6	96
3.4 Base de Datos	6	96
3.6 Aprendizaje Automático	6	96
Horas total sexto cuatrimestre		288
Cuarto Año: Séptimo Cuatrimestre		
Asignatura	hs semanales	hs totales
4.2 Práctica Profesional II	6	96
3.5 Algoritmos y Estructuras de Datos II	6	96
3.7 Electiva I	6	96
Horas total séptimo cuatrimestre		288
Cuarto Año: Octavo Cuatrimestre		
Asignatura	hs semanales	hs totales
3.8 Aprendizaje Profundo	6	96
3.9 Electiva II	6	96
4.3 Taller de Análisis e Investigación de la Práctica Profesional	6	96
Horas total octavo cuatrimestre		288

X- CONTENIDOS MÍNIMOS

A continuación, se desarrollan los contenidos mínimos de cada una de las instancias curriculares organizadas por campo formativo.

CAMPO DE FORMACIÓN COMÚN

1.1 Alfabetización digital académica

Nuevas formas de producir conocimiento en las redes. Convergencia tecnológica. Inteligencia colectiva y cultura colaborativa. Alfabetización digital y académica en la cultura digital. Web 2.0. Web 3.0. Lectura y escritura en la nube: hipertextualidad e hipermedialidad. Búsqueda de información: criterios, análisis e interpretación de fuentes de información. Abordaje crítico de las tecnologías en la cultura digital. Comunidades de aprendizaje y de práctica. Lenguaje audiovisual: producción e interpretación. Narrativas transmedia: convergencia de formatos.

1.2 Cultura general en la universidad

Formación universitaria y expansión de los horizontes culturales. Prácticas culturales. Contenidos culturales digitales. Acceso y participación en diferentes experiencias culturales (museos, reservas naturales, cine, teatro, exposiciones artísticas, exposiciones callejeras, festivales, carnavales, conciertos, centros culturales, experiencias gastronómicas, programas de radio, bibliotecas,

presentaciones de libros, talleres culturales, etc.). Análisis de las experiencias y aproximaciones a nuevas actividades.

1.3 Desafíos y problemáticas del Delta

Introducción a la vida en los Deltas. Etapas en el desarrollo local de la región. Problemáticas ambientales, sociales, económicas y culturales del Delta del Paraná. Desarrollo urbanístico y sostenibilidad. El Delta como patrimonio natural: su cuidado y conservación. Protección y preservación de los humedales del Delta. Usos, impactos y conflictos del aprovechamiento actual del Delta. El Delta y su potencialidad productiva y turística. La mejora de la calidad de vida de las poblaciones del Delta.

1.4 Inglés I

Introducción a la lectura de textos auténticos de géneros específicos de las distintas disciplinas. Estrategias de lectura para la comprensión global de textos escritos en inglés: palabras clave, transparentes, repetidas e índices tipográficos. Palabras conceptuales y estructurales. Organización textual, tema y despliegue temático. Anticipación y predicción. Elaboración del tópico del texto. Técnicas de lectura veloz: skimming y scanning. Cohesión y coherencia. Referentes contextuales: anafóricos y catafóricos; elipsis. Morfología: sufijos y prefijos. Categoría de palabras. Estructura de la información en la definición. Definición de objetos y procesos. Definiciones expandidas. El sintagma nominal. Usos del gerundio (-ing) y del participio pasado (-ed). Instrucciones. Relaciones lógicas entre proposiciones: adición, contraste, causa y efecto, enumeración. Tiempos verbales simples.

1.5 Inglés II

Estrategias de lectura para la comprensión detallada de textos pertenecientes a diversos géneros académicos y profesionales vinculados a las distintas disciplinas y carreras. Jerarquización de la información textual. Coherencia textual y avance de la información. Cadena léxica y campo semántico. Funciones retóricas: la clasificación, la descripción, la narración. El sintagma verbal; tiempo, voz y aspecto. Textos narrativos y argumentativos. Oraciones condicionales. Relaciones lógicas entre proposiciones: consecuencia, comparación, temporales, espaciales, condicionales. Tiempos verbales progresivos y perfectivos. Verbos modales simples y perfectos.

CAMPO DE FORMACIÓN BÁSICA

2.1 Introducción a la Ciencia de Datos

Obtención y carga de datos. Formato de almacenamiento de datos. Encabezados. Formato de los datos. Técnicas y herramientas de exploración y visualización de los datos. Evaluación y corrección de datos. Datos faltantes, datos incorrectos. Análisis exploratorio de los datos. Técnicas descriptivas de los datos. Muestras. Valor medio, desvío estándar. Estadísticos. Estimadores.

Concepto de correlación y covarianza. Coeficiente de correlación. Coeficiente de Pearson. Coeficiente de Spearman. Outliers.

Modelado de los datos. Modelos estadísticos básicos. Parámetros de los modelos. Regresión lineal simple. Interpretación. Nociones de ajuste por cuadrados mínimos. Estimadores de los parámetros del modelo lineal. Supuestos del modelo lineal. Diagnósticos del ajuste. Evaluación del modelo. Introducción básica a test estadísticos, residuos, QQplots. Residuos. Leverage. Outliers. Extensión de los modelos. Modelos aditivos. Regresión lineal múltiple.

2.2 Introducción al Aprendizaje Automático

Funcionamiento de algoritmos de aprendizaje automático. Validez, aplicabilidad, y limitaciones. Conjunto de datos de entrenamiento, validación y testeo. Regresión lineal. Funciones de error. Sobreajuste. Regularización. Regresión ridge y lasso. Aplicación de algoritmos de aprendizaje automático: regresión y clasificación lineal; discriminante lineal de Fischer; perceptrón; support vector machine; árboles de decisión; naive Bayes.

2.3 Análisis I

Números reales. Propiedades. Intervalos en \mathbb{R} . Módulo. Concepto de función. Dominio. Gráficas. Inyectividad y suryectividad. Función inversa. Funciones polinomiales, racionales, exponenciales y logarítmicas. Límites de funciones y Continuidad. Discontinuidades: clasificación. Cálculo diferencial. Reglas de derivación. Fórmula de Taylor. Aplicaciones de las derivadas. Intervalos de monotonía. Extremos locales. Extremos absolutos. Concavidad. Puntos de inflexión. Análisis de funciones. Gráficas aproximadas. Integral. Primitivas. Fórmula de Barrow. Cálculo integral y sus aplicaciones. Integrales impropias. Introducción a las ecuaciones diferenciales ordinarias. Métodos elementales de integración. Ecuaciones con variables separables y lineales de primer y de segundo orden.

2.4 Programación I

Lenguaje Python. Diferentes entornos de programación. Sintaxis del lenguaje. Tipos de datos básicos. Funciones y documentación. Estructuras de control básicas: condicionales, iteraciones, comprensión de listas. Estructuras de datos: diccionarios, listas, tuplas, vectores, matrices y árboles. Módulos y paquetes. Concepto de namespace. Programación orientada a objetos. Conceptos básicos, métodos, atributos, herencia. Estructuras de lectura y escritura de datos. Visualización de datos. Testeo y debuggeo de los programas. Manejo de excepciones. Control de flujos. Introducción a paquetes de cálculo y visualización. Elementos de cálculo numérico para análisis matemático.

2.5 Análisis II

Funciones vectoriales reales. Límites y continuidad. Cálculo diferencial vectorial. Regla de la cadena. Funciones implícitas. Curvas y superficies parametrizadas. Fórmula de Taylor vectorial.

Plano tangente. Problemas de máximos y mínimos. Multiplicadores de Lagrange. Campos escalares y vectoriales. Potencial. Campo tangente y normal. Formas diferenciales exactas. Análisis geométrico de ecuaciones diferenciales. Cambio de coordenadas. Coordenadas polares, esféricas y cilíndricas. Integrales múltiples. Integrales paramétricas, curvilíneas y de superficie. Aplicaciones geométricas. Operadores gradiente, divergencia, rotor y laplaciano.

2.6 Probabilidad y Estadística

Experimentos aleatorios. Frecuencias relativas. Modelo de Laplace. Probabilidad. Independencia y probabilidad condicional. Teorema de Bayes. Cálculo de probabilidades. Variables aleatorias. Funciones de densidad y de probabilidad. Funciones de distribución. Principales distribuciones discretas y continuas. Variables aleatorias vectoriales. Distribuciones conjuntas y marginales. Momentos de primer y segundo orden. Ley de Grandes Números, Teorema Central del Límite. Correlación. Nociones de procesos estocásticos. Función de autocorrelación. Procesos estacionarios. Regresión lineal. Nociones de estadística descriptiva y paramétrica. Test de hipótesis.

2.7 Estadística e Inferencia I

Estimación Puntual. Estadísticos suficientes y completos. Estimadores de momentos y de Máxima Verosimilitud. Regiones de confianza. Distribuciones multivariadas. Normal multivariada. Wishart. Regresión lineal simple y múltiple. Modelos paramétricos lineales. Stepwise regression. Regresión lineal multivariada. Modelos lineales generalizados. Teoría de la probabilidad como extensión de la lógica. Razonamiento en presencia de incertidumbre. Verosimilitud. Priors. Posteriors. Modelos con un solo parámetro. Resultados analíticos. Priors conjugados. Familia exponencial. Construcción de modelos multi-paramétricos. Modelos de mixturas finitas. Mixturas gaussianas. Responsabilidades. Algoritmo esperanza-maximización. Generalización a los modelos de variables latentes. Detección de anomalías. Validación estadística de los modelos. Distribución predictiva posterior. Validación cruzada. Consideración de la complejidad de los modelos. Nociones de teoría de la decisión. Modelos gráficos.

2.8 Programación II

Análisis de rendimiento de software. Manejo de aserciones. Prácticas en Python. Herencia múltiple. Variables privadas. Iteradores y generadores. Decoradores. Uso y funcionalidad de los decoradores. Ambientes virtuales. Paquetes de python para la ciencia de datos: pandas, seaborn. Aplicaciones. Utilización de código C, C++, y Fortran desde python. Ambientes integrales de desarrollo (IDE). Nociones vinculadas con los sistemas de control de versión. Introducción al lenguaje de análisis estadístico R.

2.9 Estadística e Inferencia II

Revisión de la regresión con modelos lineales bayesiana. Modelos generativos. Modelos de variables latentes. Responsabilidades. Generalización del algoritmo esperanza-maximización (EM). Modelos no paramétricos. Estimación de densidades. Métodos de kernel. K-means. Procesos gaussianos. Procesos de Dirichlet. Modelos jerárquicos bayesianos.

2.10 Inglés III

Escucha para la comprensión auditiva. Interacción oral. Situaciones de la vida cotidiana y propias de la actividad turística. Escritura: formato Essay Introducción, planteo del tema, cierre. Redacción de mails, notas, cartas. Vocabulario vinculado a las asignaturas del cuatrimestre. Vocabulario comercial/técnico/legal.

2.11 Inglés IV

Desempeño oral a partir de la comprensión auditiva. Interacción oral vinculada a la vida profesional/laboral. Técnica de role playing. Escritura: confección de presentaciones comerciales. Vocabulario vinculado a las asignaturas del cuatrimestre.

CAMPO DE LA FORMACIÓN DISCIPLINAR

3.1 Algoritmos y Estructuras de Datos I

Definición de algoritmos y su importancia. Características: corrección, eficiencia, complejidad. Análisis de la eficiencia temporal y espacial. Tipos de datos y su organización. Arrays. Listas enlazadas: simples, dobles y circulares. Árboles (inorden, preorden, postorden). Grafos. Listas enlazadas. Árboles binarios y árboles de búsqueda binarios (BST). Concepto, ejemplos básicos (factorial, Fibonacci). Algoritmos de ordenamiento (elección, inserción, burbuja, merge sort, quicksort). Búsqueda lineal vs. búsqueda binaria. Matrices de adyacencia. Listas de adyacencia. Caminos, ciclos, grafos dirigidos/no dirigidos. Árboles balanceados (AVL, rojo-negro). Heaps (montículos). Tablas hash. Función hash. Manejo de colisiones.

3.2 Ciencia de Datos

Series de técnicas de cálculo. Técnicas de muestreo. Muestreo simple con y sin reposición. Muestreo por rechazo. Muestreo de importancia. Importance-sampling-resampling. Gibbs sampling. Markov Chain Monte Carlo. Metropolis Hastings. Muestreo con ensambles. Extracción de features y reducción de la dimensionalidad. Variables latentes continuas. Análisis en componentes principales. PCA probabilístico. PCA bayesiano. PCA con núcleos (kernel PCA). Modelos con variables latentes no lineales. Análisis de componentes independientes. Inferencia aproximada. Métodos variacionales bayesianos. Cálculo Bayesiano Aproximado. Nociones del lenguaje de programación estadístico R y de la plataforma Stan. Visualización. Acercamientos modernos a la visualización de datos. Biplots.

3.3 Infraestructura para Ciencia de Datos

Funcionamiento de una computadora. Unidad de procesamiento central. Memoria RAM. Redes. Tiempo de latencia. Recursos compartidos. Clusters de cálculo. Cálculo usando Unidades de procesamiento de gráficos. Nociones de escala y capacidades. Conceptos básicos de cloud computing. Infraestructuras tradicionales. Almacenamiento definido por software. Soluciones de redes definidas por software. Sistemas operativos para servidores. Virtualización. Servicios de red. Servicios de directorio.

3.4 Base de Datos

Funciones de los sistemas de Base de Datos. Modelos de datos. Lenguajes de consulta. Diseño de Base de Datos. Estructuras físicas y lógicas de datos. Optimización de consultas. Transacciones. Concurrencia y recuperación. Implementaciones. Modelo de entidad-relación. Modelo relacional. SQL. Agrupamientos, subqueries, joins parciales. Sistemas de Bases de Datos. Diseño y administración de Sistemas de Bases de Datos. Escalabilidad, eficiencia y efectividad. Lenguajes de DBMS.

3.5 Algoritmos y Estructuras de Datos II

Recursión sobre listas y árboles. Programas recursivos. Tipos abstractos de datos. Tipos algebraicos: maybe, either, enumerativos, listas, árboles binarios, árboles generales. Estructuras de datos. Estructuras contenedoras: pilas, colas, diccionarios, heaps, árboles balanceados, contenedores basados en representaciones numéricas. Representación de datos en memoria. Nociones de representación e invariante de representación y su utilidad en el diseño e implementación de estructuras de datos. Uso imperativo de estructuras de datos. Iteración en listas y árboles. Modelo de memoria imperativo: stack/heap, asignación de memoria. Punteros. Variables por referencia. Listas encadenadas y sus variantes. árboles implementados con punteros. Binary heaps implementadas con arrays. Hashing. Algoritmos de recorrido, ordenamiento, búsqueda, ordenamiento y actualización. Clasificación e implementación. Complejidad Computacional. Notación O grande. Análisis teórico del tiempo de ejecución de un algoritmo Análisis práctico del tiempo de ejecución de un algoritmo. Análisis de eficiencia e implementación.

3.6 Aprendizaje Automático

Modelos de clasificación. Test set-train set. Cross-validation. Métricas para la medición de performance en regresión y clasificación. Árboles de decisión. Ensembles de árboles aleatorios (random forest). Concepto de sobreajuste. Cómo regularizar el sobreajuste. Clasificación con modelos lineales. Problemas linealmente separables. Clases balanceadas y no balanceadas. Clasificación multi-clase. Discriminante lineal de Fischer. El algoritmo del perceptrón. Método de descenso por gradiente. La función sigmoidea. Regresión logística. IRLS. Clasificadores de margen máximo. Support Vector Machines. El truco del kernel. Bases infinitas de funciones. Márgenes suaves. Combinación de modelos. Descomposición en varianza y sesgo. Bosques

aleatorio. Redes neuronales. Parametrización de las funciones de base. Funciones de activación. Redes neuronales prealimentadas. Capas ocultas y pesos. El perceptrón multicapa. Entrenamiento de redes neuronales.

3.7 Electiva I

Las materias electivas son una herramienta clave para garantizar que la Licenciatura en Ciencia de Datos se mantenga a la vanguardia de los cambios y desafíos que enfrenta la disciplina en el contexto actual. Estas asignaturas permitirán a los estudiantes profundizar en áreas de especialización según sus intereses y las demandas emergentes del sector. Las materias electivas, aprobadas por el Consejo Superior, responden a la necesidad de que la carrera evolucione y se adapte constantemente a las transformaciones y las nuevas exigencias técnicas de la Ciencia de Datos. Este enfoque flexibiliza la formación, brindando a los estudiantes la posibilidad de adquirir conocimientos específicos que complementen su formación integral, mientras se alinean con los avances y las tendencias que marcarán el futuro del uso de datos masivos..

3.8 Aprendizaje Profundo

Retropropagación. Descenso por gradiente estocástico. Gradientes que se desvanecen. Regularización. Priors gaussianos consistentes. Detención temprana. Dropout. Redes convolucionales. Invariancia de translación. Capas convolucionales y filtros. Pesos compartidos. Mapas de features. Capas de pooling. Arquitecturas de redes convolucionales. Encoders-decoders. Redes Neuronales recurrentes. Respuesta finita al impulso (FIR) y respuesta infinita al impulso (IIR). Problemas para entrenar redes recurrentes. Gradientes que se desvanecen. Long short-term memory (LSTM). Celdas LSTM. Pronóstico de series temporales univariadas y multivariadas. Utilización de redes previamente entrenadas. Transfer learning. Entrenamiento de redes profundas. Inestabilidad de los gradientes y posibles soluciones. Aplicaciones de redes recurrentes. Procesamiento del lenguaje natural. Redes stateless y statefull. Encoders-decoders para modelos de traducción. Redes no supervisadas: Autoencoders. Redes generativas antagónicas (GAN). Aplicaciones de autoencoders y GAN. Entrenamiento de GANs. Aprendizaje por refuerzo. Agente y sistema de recompensas. Aprendizaje por imitación.

3.9 Electiva II

Las materias electivas son una herramienta clave para garantizar que la Licenciatura en Ciencia de Datos se mantenga a la vanguardia de los cambios y desafíos que enfrenta la disciplina en el contexto actual. Estas asignaturas permitirán a los estudiantes profundizar en áreas de especialización según sus intereses y las demandas emergentes del sector. Las materias electivas, aprobadas por el Consejo Superior, responden a la necesidad de que la carrera evolucione y se adapte constantemente a las transformaciones y las nuevas exigencias técnicas de la Ciencia de Datos. Este enfoque flexibiliza la formación, brindando a los estudiantes la

posibilidad de adquirir conocimientos específicos que complementen su formación integral, mientras se alinean con los avances y las tendencias que marcarán el futuro del uso de datos masivos.

4. CAMPO DE LA PRÁCTICA PROFESIONAL

4.1 Práctica Profesional I y 4.2 Práctica Profesional

Se orienta a la aplicación y desarrollo de conocimientos en una práctica específica, inserta en el contexto institucional local-regional tanto extrauniversitario como intrauniversitario, favoreciendo la gradualidad en el desarrollo de competencias y la diversidad de experiencias situacionales en el mundo del trabajo.

Los objetivos de estas prácticas son:

- Vincular al estudiante con la dinámica de un ámbito laboral específico.
- Fortalecer las competencias expresivas e instrumentales fundamentales para el desempeño profesional.
- Desarrollar capacidades para la observación y análisis de situaciones desde una perspectiva integral.
- Fortalecer las capacidades de utilización de tecnologías en el mundo del trabajo contemporáneo.
- Promover la inserción profesional del alumno en instituciones locales y regionales.

4.3 Taller de Análisis e Investigación de la Práctica Profesional

Tiene como objetivo proporcionar un espacio donde los estudiantes reflexionen sobre su experiencia de práctica profesional, desarrollen habilidades de análisis crítico y capacidad para investigar las realidades del sector en que se desempeñan. Este taller busca potenciar las competencias de investigación, análisis y presentación de proyectos relacionados con las prácticas profesionales, con el fin de promover un aprendizaje autónomo y una mejor comprensión de las dinámicas laborales. A través de este taller, los estudiantes integrarán sus conocimientos y experiencias en proyectos prácticos, propondrán soluciones innovadoras a problemas reales y profundizarán en la evaluación del impacto de las prácticas realizadas en el contexto laboral y en la sostenibilidad del sector.