Ciencia de Datos

Planificación Ciclo lectivo 2026

1. Datos administrativos de la asignatura

Departamento:	Sistemas de Información	Carrera	Ing. en Sistemas de Información
Asignatura:	**Ciencia de Datos**		
Nivel de la carrera	5to Año	Duración	Cuatrimestral
Plan	2023 Ord. 1877	Modalidad	Presencial
Bloque Curricular:	Tecnologías Aplicadas		
Área	Sistemas Inteligentes	RTF	6
Horas cátedra semanal:	3	Horas reloj anual total:	72
Carga horaria no presencial semanal	----------	% horas no presenciales	----------
Profesor/es Titular/Asociado/Adjunto:		Dedicación:	1 DE
Auxiliar/es de 1°/JTP		Dedicación:	1 DS
Competencias	Específicas	Genéricas	
	CE1.1 – CE4.1 – CE5.1	CG3 – CG4 – CG6	

2. Presentación, Fundamentación

Las grandes transformaciones y los innumerables avances tecnológicos, científicos, y sociales, que se han generado en las últimas décadas han posibilitado acceder y obtener conocimientos de diferentes maneras. En la actualidad las organizaciones desean sacar provecho de los datos almacenados en múltiples dispositivos, sistemas y sobre todo los datos alojados en la nube.

Los datos generados y guardados en diferentes fuentes se multiplican día a día debido, en gran medida, al avance en el desarrollo, al poder de cálculo, las reducciones del costo del hardware, la diseminación de nuevas aplicaciones de software, entre otros. Un impulsor fundamental de este crecimiento ha sido internet, mediado por la masificación de las aplicaciones de colaboración, los sistemas transaccionales de gestión en la nube, el intercambio de email y redes sociales, entre muchos otros.

En este contexto, la ciencia de datos se ha consolidado como una disciplina para descubrir conocimiento basándose en los datos accesibles en los entornos más diversos. La ciencia de datos ofrece herramientas y metodologías para tratar los problemas de obtención, depuración y aplicación de modelos para análisis, la transformación de datos en información y la información en conocimiento para la toma de decisiones

Relación de la asignatura con el perfil de egreso.

En relación con el perfil de egreso, al ser una materia con gran carga técnica aporta sobre todo al desarrollo de innovaciones tecnológicas para la mejora de calidad de vida de la sociedad. Se impulsan competencias para que el estudiante capacitado asuma un liderazgo efectivo en la gestión técnica y metodológica.

Relación de la asignatura con los alcances del título.

En relación con los alcances del título de Ingeniería en Sistemas de Información esta materia aporta al alcance 7 en lo que respecta a especificar, proyectar, desarrollar, implementar y evaluar modelos relacionados a la inteligencia artificial y la ciencia de datos.

3. Relación de la asignatura con las competencias de egreso de la carrera

Se detalla en la siguiente tabla, la relación de la asignatura con las competencias de egreso específicas, genéricas tecnológicas, sociales, políticas y actitudinales de la carrera. Indicando a cuáles competencias de egreso tributa y en qué nivel (No Tributa, 1=bajo, 2=medio, 3=alto).

Competencias	Nivel
Competencias genéricas tecnológicas (CG):	
CG. 1. Identificar, formular y resolver problemas de ingeniería.,	No Tributa
CG. 2. Concebir, diseñar y desarrollar proyectos de ingeniería.	No Tributa
CG. 3. Gestionar, planificar, ejecutar y controlar proyectos de ingeniería.	3
CG. 4. Utilizar de manera efectiva las técnicas y herramientas de aplicación en la ingeniería.	3
CG. 5. Contribuir a la generación de desarrollos tecnológicos y/o innovaciones tecnológicas.	No Tributa
Competencias Genéricas Sociales Políticas y Actitudinales (CG):	
CG. 6. Desempeñarse de manera efectiva en equipos de trabajo.	3
CG. 7. Comunicarse con efectividad	No Tributa
CG. 8. Actuar con ética, responsabilidad profesional y compromiso social, considerando el impacto económico, social y ambiental de su actividad en el contexto local y global.	No Tributa
CG. 9. Aprender en forma continua y autónoma.	No Tributa
CG. 10. Actuar con espíritu emprendedor	No Tributa
Competencias Específicas (CE):	
CE. 1.1. Especificar, proyectar y desarrollar sistemas de información para concebir soluciones tecnológicas que permitan resolver situaciones en las organizaciones mediante el empleo de metodologías de sistemas y tecnologías asociadas a los sistemas de información	3
CE. 1.2. Especificar, proyectar y desarrollar sistemas de comunicación de datos, evaluando posibles soluciones tecnológicas disponibles para dar soporte a los sistemas de información en lo referido al procesamiento y comunicación de datos.	No Tributa
CE. 1.3. Especificar, proyectar y desarrollar software para la elaboración de soluciones informáticas con el propósito de resolver problemas estratégicos y operativos, así como de servicios y de negocios, en el marco de una actividad económica que sea social y ambientalmente sustentable.	No Tributa
CE. 2.1. Proyectar y dirigir lo referido a seguridad informática para seleccionar y aplicar técnicas, herramientas, métodos y normas, garantizando la seguridad y privacidad de la información procesada y generada por los sistemas de información.	No Tributa

CE. 3.1. Establecer métricas y normas de calidad de software para medir, evaluar, controlar y monitorear el rendimiento, impulsando mejoras de acuerdo a técnicas y normas vigentes definidas por los organismos de estandarización.	No Tributa
CE. 4.1. Certificar el funcionamiento, condición de uso o estado de sistemas de información, sistemas de comunicación de datos, software, seguridad informática y calidad de software para asegurar la generación de los resultados deseados en función de restricciones de tiempo y recursos establecidos.	3
CE. 5.1. Dirigir y controlar la implementación, operación y mantenimiento de sistemas de información, sistemas de comunicación de datos, software, seguridad informática y calidad de software, a los fines de alcanzar los objetivos fijados por la organización.	3
CE. 6.1. Asesorar y capacitar a organizaciones, empresas, organismos públicos o privados en la adquisición, instalación y uso, en lo que respecta a sistemas de información, sistemas de comunicación de datos, software, seguridad informática y calidad de software, a los fines de un uso correcto de los sistemas intervinientes.	No Tributa
CE. 7.1. Realizar pericias, tasaciones y arbitrajes relacionados con su actividad profesional, respetando marcos normativos y jurídicos con el objeto de asesorar a las partes o a los tribunales de Justicia.	No Tributa

4. Propósito

Obtener las capacidades de gestionar un proyecto de ciencias de datos utilizando estrategias de extracción y transformación datos para dar solución a situaciones problemáticas en contexto donde ejecuten su profesión.

5. Objetivos establecidos en el Diseño Curricular

- Distinguir estrategias para gestionar un proyecto de ciencia de datos.
- Aplicar estrategias de extracción, visualización y transformación de datos en forma de patrones útiles y aplicables en el desarrollo de sistemas inteligentes.
- Crear modelos que resuelvan situaciones problemáticas particulares en el contexto del negocio

6. Contenidos Mínimos establecidos en el Diseño Curricular

- Gestión de proyectos de ciencia de datos.
- Análisis exploratorio de los datos.
- Visualización de datos.

- Extracción y Transformación de datos.
- Ingeniería de datos.
- Algoritmos de Clasificación y Agrupamiento.
- Minería de Texto.

7. Resultados de aprendizaje

Se describen los Resultados de aprendizaje (RA) a promover en el desarrollo de la asignatura.

Identificador del RA	Redacción
RA1	[Gestiona] de manera integral [un proyecto de ciencia de datos tomando en cuenta los diferentes abordajes metodológicos, las herramientas disponibles y las limitaciones técnicas] para [desarrollar soluciones que apoyen la toma de decisiones]
RA2	[Desarrolla] [un proceso de extracción, transformación y análisis exploratorio] para [adecuar los datos según los requerimientos de los modelos a utilizar], considerando [las herramientas seleccionadas]
RA3	[Evalúa] [herramientas y técnicas de visualización] para [analizar e interpretar los datos obtenidos de una manera comprensible]
RA4	[Implementa] [herramientas estadísticas, matemáticas, algorítmicas y tecnológicas] para [extraer valor de los datos en la forma de patrones útiles y aplicables para aportar nuevo conocimiento].

8. Asignaturas correlativas previas

Para cursar y rendir

Cursada:

- Simulación

Aprobada:

- Probabilidad y Estadística.
- Base de Datos

9. Asignaturas correlativas posteriores

No aplica.

10. Programa analítico, Contenidos temáticos

Ejes temáticos	Contenidos
Gestión de proyectos de ciencia de datos	Ciencia de datos El rol del científico de datos Gestión de proyectos en ciencia de datos Landscape de la ciencia de datos Herramientas para la ciencia de datos
Extracción, transformación y carga de datos (ETL)	Técnicas de extracción de datos Tipo de fuentes de datos Conexiones, drivers, Apis Técnicas de transformación de datos Limpieza de datos Análisis exploratorio Calidad de datos Almacenes de datos y técnicas de carga Arquitecturas. Tipos de almacenes. Modelos. Landscape de herramientas (almacenes). Procesamientos discretos y continuos. Procesamiento batch o Sistemas de colas Procesamiento continuo, streaming o en tiempo real.
Visualización y análisis exploratorio de datos	Tipos de gráficos Análisis exploratorio mediante visualización Técnicas para la visualización de datos
Algoritmos de Clasificación, agrupamiento y minería de Texto	Aprendizaje supervisado Regresión Lineal Regresión Polinomial Ajuste y Selección del Modelo Regresión Logística Árboles de Decisión Random Forest Aprendizaje no supervisado Clustering Jerárquico Clustering Particional Minería de texto

Cuadro de Carga Horaria de Formación Práctica

Tipo de Formación Práctica	Carga horaria	
	Presencial	A Distancia
Instancias supervisadas de formación práctica (prácticas en diferentes ámbitos tales como aula, laboratorios, campo u otros)	32 Hs	
Proyecto Integrador		
Práctica profesional supervisada		
Carga horaria total	32 Hs	

11. Metodología de enseñanza

La metodología de enseñanza se abordará de la siguiente forma:

1- Los correspondientes al manejo de proyectos, proceso de extracción/transformación y carga y visualización y análisis de datos.

Centrado en la metodología de aprendizaje basado en proyectos, donde los estudiantes son los protagonistas de su propio proceso de aprendizaje promoviendo la investigación, sistematización y aplicación del conocimiento aplicado al proyecto (trabajo práctico integrador). El docente asume el papel de guiar a los estudiantes durante su jornada de aprendizaje, aplicando tareas, dinámicas y actividades en formatos de casos de estudio, ejercicios, investigaciones y sobre todo su experiencia profesional. Se utilizan las clases especialmente para el debate e intercambio de informaciones, posibilitando que los propios estudiantes aporten conocimientos adquiridos al colectivo de la cátedra.

2- El manejo de algoritmos y minería de texto.

La formación práctica ocupará un lugar importante dentro del desarrollo de esta asignatura. Debido a la complejidad e importancia que tiene el manejo de los algoritmos y la minería de datos para la materia, se utiliza una estrategia pedagógica en línea con la metodología de formación experimental en laboratorio y estudios de casos. Se promueve que además de aprender cada algoritmo, los estudiantes entiendan cómo es su estructura interna y su funcionamiento para un correcto uso. De esta forma se desarrollan actividades prácticas con el educador, ejercicios, pruebas y desafíos para que los alumnos no solo reciban información, sino que también puedan aplicar lo que aprendieron en clase de forma adecuada en el trabajo práctico integrador. A partir de las actividades los estudiantes pueden analizar, interpretar, argumentar, explicar los resultados y/u observaciones y debatir con sus pares, para lo cual deben relacionar saberes

previos y los alcanzados en el proceso alterando las diferentes variables que proponen los algoritmos.

Esta estrategia también se complementará con material complementario que pueda ayudar a ampliar los conocimientos y profundizar sobre los algoritmos

Resumen de Competencias - Resultados de Aprendizaje - Metodologías de Enseñanza de la asignatura.

COMPETENCIA - NIVEL DE TRIBUTACIÓN	RESULTADO DE APRENDIZAJE (RA)
CE1.1(3) – CE4.1(3) – CE5.1(3) - CG3(3) – CG4(3) – CG6(3)	RA1 - RA2 - RA3 - RA4

EJES TEMÁTICOS - CONTENIDOS

Gestión de proyectos de ciencia de datos - Extracción, transformación y carga de datos (ETL) - Visualización y análisis exploratorio de datos - Algoritmos de Clasificación, agrupamiento y minería de Texto

ACTIVIDADES DEL ESTUDIANTE (FORMATIVAS)

El **estudiante** deberá realizar las siguientes actividades para alcanzar los **RA** propuestos:

- Lectura y análisis de material bibliográfico.
- Trabajo Práctico Integrador
- Trabajo Práctico Grupal
- Cuestionarios.
- Resolución de Estudio de Casos
- Instancia de Evaluación Individual
- Resolución de problemas
- Exposición oral

ESTRATEGIAS DE ENSEÑANZA

El **docente** diseñará las siguientes **actividades pedagógicas**:

- Clase expositiva dialogada
- Taller / Seminario.
- Acompañamiento continuo.
- Diseño de problemas/ejercicios.
- Diseño de Casos de Estudios.
- Diseño Instancia de Evaluación.
- Retroalimentación formativa.
- Coordinación de trabajo colaborativo.

12. Recomendaciones para el estudio

- Conocer previamente el programa analítico de la materia.
- Revisar los contenidos requeridos de materias previas como probabilidad y estadística, simulación, investigación operativa e inteligencia artificial.
- Apropiarse del proceso de enseñanza aprendizaje y lograr su máximo potencial, de forma individual y como colectivo.
- Durante la cursada, ampliar los conocimientos mediante todos los recursos disponibles.
- Responsabilidad y colaboración con el equipo de trabajo para lograr los mejores resultados y desarrollar las competencias.

13. Metodología de evaluación

Los métodos de evaluación serán diversos, realizando un monitoreo de las actividades propuestas a los estudiantes fuera del aula, y analizando la participación en ejercicios durante las clases.

También se utilizarán recursos interactivos en tiempo real con incentivos reales sobre los efectos de la evaluación individual del estudiante, generando así una doble recompensa de efecto directo de su condición evaluativa y una a largo plazo de conocimiento aplicado, basado en actividades recreativas y de realización grupal.

Finalmente se realizarán tres evaluaciones parciales y directas de contenidos que permitan tener revisiones periódicas de las competencias esperadas.

Evaluación de cada resultado de aprendizaje

Para las actividades propuestas dentro y fuera del contexto del aula se evaluará el desarrollo y la participación activa en las mismas, sirviendo como un sensor general y sensor del desarrollo de las competencias en el estudiante. No se tomará en cuenta para esta etapa evaluativa el nivel y la calidad de lo realizado, si no el interés y la involucración en la asignatura.

Las actividades de tiempo real grupales tendrán una retroalimentación inmediata de los resultados y serán claros en los efectos evaluativos del alumno.

Las 8 evaluaciones parciales serán revisadas por el equipo docente y se hará una devolución numérica y retroalimentación de los contenidos que el estudiante desarrolle.

Condiciones de aprobación

- Asistencia del 75%
- Desarrollo de las actividades propuestas extra-áulicas.
- Desarrollo de las actividades propuestas en el aula.
- Realización de al menos el 80% de los cuestionarios de evaluación.

- Participación de las actividades en tiempo real y grupales.
- La nota se obtendrá en base a los cuestionarios realizados en clase y los trabajos grupales integradores que asignen créditos.
- Obtener nota mayor a seis.

Se considera en condición de ALUMNO REGULAR a quien cumpla las condiciones del punto anterior, u obtenga nota mayor a seis en la evaluación integradora final.

Se considera en condición de APROBACIÓN DIRECTA a quienes obtengan una nota mayor a ocho y cumplan además las condiciones de ALUMNO REGULAR, u obtenga nota mayor a ocho en la evaluación integradora final y cumpla con las condiciones de asistencia.

Aquellos estudiantes que no cumplan las condiciones anteriores se consideran en condición de LIBRES, debiendo recursar la asignatura.

14. Cronograma de clases año 2026

Semana / Fecha	Contenido / Actividad
Semana 1	Metodologías (CRISP-DM, TDSP). Definición de problemas de negocio vs. problemas de datos. Caso de estudio: Diseñar la hoja de ruta de un proyecto real desde la captura hasta el despliegue.
Semana 2	Arquitecturas de datos (Data Lakes vs. Data Warehouses). Configuración de un entorno de trabajo (Docker/Conda) y conexión a fuentes de datos externas.
Semana 3	Pipelines de datos. Limpieza de datos (Data Wrangling). Integración de fuentes heterogéneas. Ejercicio práctico de depuración de un dataset ruidoso usando Python/Pandas.
Semana 4	Creación de variables, normalización, escalado y manejo de valores faltantes (imputación). Laboratorio sobre selección de características y reducción de dimensionalidad básica.
Semana 5	Estadística descriptiva aplicada. Identificación de sesgos, outliers y correlaciones. Reporte de diagnóstico sobre un dataset desconocido.
Semana 6	Teoría de la percepción visual. Tipos de gráficos según el dato. Narrativa de datos (Storytelling). Creación de un dashboard interactivo utilizando bibliotecas como Plotly, Seaborn o herramientas como Metabase.
Semana 7	Presentación del pipeline de datos y el análisis exploratorio del proyecto final.

Semana 8	Regresión Logística y K-Nearest Neighbors (KNN). Métricas de evaluación (Precisión, Recall, F1-Score). Implementación de un clasificador binario y análisis de la Matriz de Confusión.
Semana 9	Árboles de Decisión y Random Forest. Conceptos de Overfitting y Underfitting. Comparativa de modelos: ¿Por qué un modelo complejo no siempre es el mejor?
Semana 10	Aprendizaje no supervisado. K-Means, DBScan y algoritmos jerárquicos. Segmentación de clientes basada en comportamiento de compra.
Semana 11	Cross-validation, búsqueda de hiperparámetros (GridSearch). Optimización del modelo de clasificación desarrollado en la semana 9.
Semana 12	Preprocesamiento de texto (Tokenización, Stopwords, Stemming/Lemmatization). Representación (Bag of Words, TF-IDF). Análisis de frecuencia de palabras en un corpus de noticias.
Semana 13	Análisis de sentimiento y modelado de tópicos (LDA). Clasificador de reseñas de productos (Sentimiento positivo/negativo).
Semana 14	Sesgos algorítmicos. Privacidad (GDPR). Introducción simple a MLOps (cómo vive un modelo en el mundo real). Debate sobre casos éticos.
Semana 15	Mentoría individual y resolución de problemas técnicos en los proyectos de los estudiantes.
Semana 16	Exposición de los proyectos integradores: desde la extracción de los datos hasta la visualización y predicción final.

15. Recursos necesarios

- Aula o laboratorio de computación.
- Disponibilidad de infraestructura y equipamiento. Los espacios físicos y equipamiento están disponibles para el desarrollo de las actividades prácticas. Eventualmente, los alumnos pueden realizar las actividades prácticas a través del uso de máquinas virtuales en sus notebooks personales dentro del aula sin necesidad de la utilización de laboratorios.

16. Referencias bibliográficas (citadas según Normas APA)

Bibliografía Obligatoria:

- Wes McKinney (2017). Python for Data Analysis. O'Reilly Media.
- Joe Reis & Matt Housley (2022). Fundamentals of Data Engineering. O'Reilly Media.
- Giuseppe Bonaccorso (2018). Machine Learning Algorithms Second Edition. Packt Publishing Ltd.

Bibliografía complementaria:

- Joel Grus (2015). Data Science From Scratch : First Principles With Python. O'Reilly Media.
- John Paul Mueller & Luca Massaron (2019). Python for Data Science 2º Edition. John Wiley & Sons.
- Lilian Pierson(2021). Data Science 3º Edition. John Wiley & Sons.

17. Función Docencia

Dictado de las clases
Elaboración de plan de trabajos de situaciones prácticas
Disponibilidad en tiempo y forma de recursos materiales (infraestructura y equipamiento).
Actualización de material bibliográfico
Aprobación de instrumentos de evaluación.
Dictado de las clases colaborativas en resolución de problemas.

18. Reuniones de asignatura y área

Reuniones periódicas de asignatura y se prevén dos reuniones de área en el año.

19. Atención y orientación a las y los estudiantes

Las clases de consultas se coordinan con los estudiantes según sus necesidades.
Se enfatiza el uso de los recursos de comunicación asincrónica y sincrónica que proporciona la plataforma virtual (Moodle).
Se planifican encuentros individuales o grupales por medio de videoconferencia por las diversas plataformas.