UNIVERSIDAD NACIONAL DE ASUNCIÓN
FACULTAD POLITÉCNICA
CARRERA INGENIERÍA EN INFORMÁTICA
PLAN 2008
PROGRAMA DE ESTUDIO

Resolución 25/07/09-00 Acta 1215/07/04/2025
ANEXO 05

I. IDENTIFICACIÓN

1. Asignatura : Electiva – Data Science
2. Semestre : Según la alternativa en la que se inscribe el estudiante
3. Horas semanales : 7 horas
   - 3.1. Clases teóricas : 3 horas
   - 3.2. Clases prácticas : 4 horas
4. Total real de horas disponibles : 112 horas
   - 4.1. Clases teorías : 48 horas
   - 4.2. Clases prácticas : 64 horas

II. JUSTIFICACIÓN

A medida que nuestra economía, la sociedad y la vida diaria se vuelven cada vez más dependientes de los datos, los nuevos graduados universitarios que ingresan a la fuerza laboral deben tener las habilidades para analizar los datos de manera efectiva. Así, la ciencia de datos a emergido como un campo multifacético utilizado para obtener información a partir de datos complejos.
La Ciencia de datos se puede definir como una nueva "trans-disciplina" que se consolida de la convergencia mediante la síntesis de un relevante número de disciplinas y conocimientos incluidos la computación, las estadísticas, la misma informática, comunicaciones, administración y hasta la sociología para estudiar cómo la enorme cantidad de datos hoy disponibles, pueden convertirse en nuevos conocimientos para dominios bien específicos como, la medicina, los negocios, las ciencias políticas y cualquier otra área en la que se genera gran cantidad de datos que hoy no son suficientemente explotados [Longbing Cao, 2017].
Adquirir habilidades proporcionadas por la ciencia de datos es esencial para los estudiantes universitarios ahora y en el futuro. La evolución emergente de la ciencia de datos significa oportunidades para la investigación de vanguardia, la innovación tecnológica y una nueva economía de datos. Si se trazan paralelos entre la evolución de Internet y la evolución de la ciencia de datos, el futuro y el impacto socioeconómico y cultural de la ciencia de datos no tendrán precedentes, aunque todavía no se pueden cuantificar.

III. OBJETIVOS

3.1. Formar profesionales con competencias destacadas en investigación y aplicación en el área de Ciencia de datos que estén capacitados para abordar nuevos desafíos en el manejo de la información en diversos dominios de conocimiento.

IV. PRE – REQUISITO

Para cursar esta asignatura, que se ofrece en la carrera como Electiva, el estudiante deberá cumplir con los prerrequisitos establecidos según la electiva en la que se inscribe, de acuerdo con la siguiente tabla:

Alternativa	Porcentaje de créditos aprobados	Cantidad de créditos requeridos
Electiva 1	55%	184
Electiva 2, Electiva 3, Electiva 4, Electiva 5	70%	235
Electiva 6, Electiva 7	80%	268

V. CONTENIDO

5.1. Unidades programáticas

5.1.1. Introducción a la Ciencia de datos
5.1.2. Programación con Python para Ciencia de datos
5.1.3. Análisis exploratorio de datos
5.1.4. Ingeniería de datos
5.1.5. Machine Learning
5.1.6. Visualización de datos y comunicación de los resultados
5.1.7. Recuperación de la información desde repositorios de datos masivos
5.1.8. Cuestiones Legales en el manejo de los datos
5.1.9. Presentación de un Proyecto Final

5.2. Desarrollo de las unidades programáticas

5.2.1. Introducción a la Ciencia de datos

Se abordará una visión general completa de los temas centrales en Ciencia de Datos, tanto desde el punto de vista matemático como desde la perspectiva de la Inteligencia Artificial. Se pondrá especial énfasis en: las clases básicas de técnicas y métodos, los fundamentos teóricos de la ciencia de datos, la inteligencia computacional, y los principales dominios de aplicación de la ingeniería del conocimiento.

5.2.2. Programación con Python para Ciencia de datos

Conceptos de programación y aplicaciones usando el lenguaje Python. Introducción a las bibliotecas esenciales para Ciencia de datos

5.2.3. Análisis exploratorio de datos

Técnicas exploratorias esenciales para la comprensión de los datos. Estas técnicas generalmente se aplican antes de que comience el modelado formal y pueden ayudar en la percepción de los modelos estadísticos más complejos. Las técnicas exploratorias también son importantes para eliminar o agudizar las hipótesis potenciales que pueden ser abordadas por los datos.

5.2.4. Ingeniería de datos

Manejo de los datos crudos: Técnicas de depuración, limpieza y preparación.

5.2.5. Machine Learning

Técnicas básicas y avanzadas de aprendizaje automático. Se enfatizará el uso práctico de las técnicas presentadas y los problemas para desarrollar aplicaciones reales de aprendizaje automático. El curso está acompañado de prácticas de laboratorio que ayudan al alumno a comprender el funcionamiento de los algoritmos de aprendizaje automático.

5.2.6. Visualización de datos y comunicación de los resultados

Exploración de datos en forma visual y construcción de un relato para comunicar los resultados obtenidos del análisis de los datos.

5.2.7. Recuperación de la información desde repositorios de datos masivos

Fundamentos de la recuperación de información, su importancia, así como casos prácticos.

5.2.8. Cuestiones Legales en el manejo de los datos

Tomar conciencia de la normativa que se aplica dentro del ámbito de la Ciencia de Datos tanto a nivel nacional como internacional, los límites que se imponen y las obligaciones que conlleva el manejo de datos masivos.

5.2.9. Presentación de un Proyecto Final

VI. ESTRATEGIAS METODOLÓGICAS

6.1. Exposición de fundamentos teóricos apoyados por medios audiovisuales.
6.2. Trabajos de investigación y evaluación de herramientas y técnicas.
6.3. Diseño y Programación de un Proyecto de Big Data
6.4. Trabajos en Laboratorio.

VII. MEDIOS AUXILIARES

7.1. Pizarras acrílicas
7.2. Marcadores
7.3. Borrador de pizarra acrílica
7.4. Computadores
7.5. Proyector multimedia
7.6. Plataforma Virtual EDUCA
7.7. Sala de laboratório equipada para las prácticas
7.8. Acceso a internet

VIII. EVALUACIÓN

La evaluación sobre el aprendizaje y conocimiento adquiridos por el alumno se realizará de acuerdo a lo establecido en el reglamento de la Facultad Politécnica de la UNA.

IX. BIBLIOGRAFÍA

- Baumer, Ben. "A data science course for undergraduates: Thinking with data." *The American Statistician* 69.4 (2015): 334-342.
- Cao, Longbing. *"Data science: challenges and directions." Communications of the ACM* 60.8 (2017): 59-68.
- Cover, Thomas M., and Joy A. Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

- Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction.* Springer Science & Business Media, 2009
- Kelleher, John D., and Brendan Tierney. *Data science.* MIT Press, 2018.
- Larose, Daniel T. *Data mining and predictive analytics.* John Wiley & Sons, 2015.
- VanderPlas, Jake. *Python data science handbook: essential tools for working with data.* " O'Reilly Media, Inc.", 2016.

Enlaces Web

- https://www.python.org/
- https://www.ibm.com/developerworks/library/ba-intro-data-science-1/ba-intro-data-science-1-pdf.pdf