



Campus de la UNA
SAN LORENZO-PARAGUAY

UNIVERSIDAD NACIONAL DE ASUNCIÓN
FACULTAD POLITÉCNICA
CONSEJO DIRECTIVO

RESOLUCIÓN 25/18/160-00
ACTA 1226/25/08/2025

“POR LA CUAL SE APRUEBA EL PROGRAMA DE ESTUDIO DE LA ASIGNATURA TÉCNICAS DE EXTRACCIÓN DE DATOS, DE CARRERAS DE GRADO, SEDE SAN LORENZO, FILIAL VILLARRICA Y FILIAL CORONEL OVIEDO”

VISTO: El Memorando DA/1707/2025 del Director Académico de la FP-UNA, Prof. MSc. Felipe Santiago Uzabal Ecurra, con el cual remite el Memorando CCPTCC/030/2025 de la Comisión Coordinadora del Proyecto de Transformación Curricular de Carreras de Grado de la FP-UNA, en el que presenta la propuesta de Programa de Estudio de Asignatura de las Carreras de Grado.

CONSIDERANDO: La Ley 4995/2013 de Educación Superior, el Estatuto de la Universidad Nacional de Asunción y las deliberaciones sobre el tema.

Que la Comisión Coordinadora del Proyecto de Transformación Curricular de Carreras de Grado, solicita la aprobación del Programa de Estudio de la asignatura **“Técnicas de Extracción de Datos”**, la cual es común entre Carreras de Grado.

Que los programas fueron elaborados conforme a las disposiciones establecidas por el Consejo Nacional de Educación Superior (CONES) en materia de **créditos académicos**, según lo dispuesto en la Resolución CONES N.º 221/2024, que regula el *Sistema de Créditos Académicos – Paraguay* y los criterios para su publicación en las carreras de grado.

**EL CONSEJO DIRECTIVO DE LA FACULTAD POLITÉCNICA
RESUELVE:**

25/18/160-01 APROBAR el Programa de Estudio de la Asignatura **“Técnicas de Extracción de Datos”**, de carreras de Grado, Sede San Lorenzo, Filial Villarrica y Filial Coronel Oviedo, detallado en el ANEXO 122 de la presente Acta.

25/18/160-02 COMUNICAR, copiar y archivar.

Prof. Abg. Joel Arsenio Benítez Santacruz
Secretario



Prof. Ing. Silvia Teresa Leiva León, MSc.
Presidenta



Campus de la UNA
SAN LORENZO-PARAGUAY

UNIVERSIDAD NACIONAL DE ASUNCIÓN
FACULTAD POLITÉCNICA
CONSEJO DIRECTIVO

Resolución 25/18/160-00 Acta 1226/25/08/2025
ANEXO 122

DEPARTAMENTO DE ENSEÑANZA DE INFORMÁTICA
PROGRAMA DE ESTUDIO

I. IDENTIFICACIÓN

Nivel				Grado									
Asignatura				Técnicas de Extracción de Datos									
Carrera				Plan		Sede/Filial		Carácter		Semestre		Prerrequisitos	
Ingeniería en Informática				2026		Sede San Lorenzo		Electiva 1		Séptimo		Seguridad Informática, Programación de Aplicaciones 2, Bases de Datos 2, Ingeniería de Software 1.	
Licenciatura en Ciencias Informáticas				2026		Sede San Lorenzo/Filial Villarrica/Filial Coronel Oviedo		Electiva		***		Ingeniería de Software I, Base de Datos II, Seguridad Informática, Programación Frontend.	
Semanal					Periodo								
HT		HP		HTD	HTI	HS	PA	THTD		THTI		THA	CA-PY
2		2		4	4	8	18	72		72		144	5

- *HT: Horas Teóricas semanales.
- *HP: Horas Prácticas semanales.
- *HTD: Horas semanales de Trabajo académico con acompañamiento Docente.
- *HTI: Horas semanales de Trabajo académico Independiente del estudiante.
- *HS: Horas Semanales (HTD+HTI).
- *PA: Periodo Académico en semanas.
- *THTD: Total de Horas de Trabajo académico con acompañamiento Docente (HTD*PA).
- *THTI: Total de Horas de Trabajo académico Independiente del estudiante (HTI*PA).
- *THA: Total de Horas de trabajo Académico (THTD+THTI).
- *CA-PY: Créditos académicos de la asignatura.

II. FUNDAMENTACIÓN

La extracción de datos es una etapa crítica en el proceso de análisis de datos y minería de datos. Esta asignatura proporciona a los estudiantes una comprensión profunda de las técnicas y herramientas utilizadas para extraer, limpiar, transformar y cargar datos desde diversas fuentes. Estas habilidades son esenciales en el mundo actual, donde el manejo eficaz de grandes volúmenes de datos es crucial para la toma de decisiones empresariales, la investigación científica y el desarrollo tecnológico.

El objetivo principal es el de consolidar en el estudiante conocimientos de técnicas y herramientas para la extracción, transformación y carga de datos aplicable a las áreas de análisis de datos, inteligencia de negocios e inteligencia artificial. Se pretende que el alumno conozca metodologías de extracción de datos aplicando técnicas computacionales que podrán ser empleadas en el abordaje de problemas de investigación o de su actividad profesional.

En relación a la naturaleza de la asignatura, se aborda de manera teórico-práctica, se combinarán conceptos teóricos con ejercicios prácticos. La organización de la asignatura se basa en los ejes temáticos,

temáticos, se incluyen conceptos fundamentales como: Técnicas de web scraping, APIs, limpieza y almacenamiento de datos extraídos.

III. COMPETENCIAS DEL PERFIL DE EGRESO ASOCIADAS

- 1. Liderar y trabajar en equipo con eficacia y responsabilidad tomando decisiones basadas en evidencias.
- 2. Evaluar el comportamiento de diversos fenómenos disciplinares e interdisciplinares relacionados con la informática con una visión de sistema, mediante modelos teóricos validados y actualizados, capaces de abarcarlos integralmente en un contexto de incertidumbre.
- 3. Seleccionar, utilizar y construir instrumentos innovadores asociados al ejercicio de la informática.
- 4. Aplicar, producir y difundir conocimientos técnicos y científicos en el área de la informática.
- 5. Planificar, proyectar, diseñar y ejecutar proyectos sostenibles e integrales para la resolución de problemas, la mejora y la innovación en el ámbito de la informática.
- 6. Interpretar, modelar y comunicar información referida a la informática en forma gráfica.

IV. ORGANIZACIÓN DE LA ASIGNATURA

Unidades	Contenidos	Resultados de aprendizaje
1. Introducción a la Extracción de Datos.	1.1. Conceptos básicos de extracción de datos. 1.2. Herramientas y técnicas de extracción de datos. 1.3. Fuentes de datos: estructuradas, semiestructuradas y no estructuradas. 1.4. Introducción a los procesos ETL..	1. Comprende el concepto de extracción de datos y sus implicancias técnicas. 2. Reconoce la evolución histórica de las herramientas y técnicas de extracción de datos. 3. Reconoce la evolución histórica de las fuentes de datos tanto estructuradas como no estructuradas. 4. Comprende la diferencia entre fuentes de datos estructuradas y no estructuradas. 5. Reconoce y comprende procesos de ETL.
2. Técnicas de Extracción desde Bases de Datos Relacionales y NoSQL.	2.1. Extracción desde bases de datos relacionales usando SQL. 2.2. Extracción desde bases de datos NoSQL (MongoDB, Cassandra, etc.). 2.3. Optimización de consultas para extracción de datos..	1. Modela un problema de extracción de datos desde bases de datos estructuradas. 2. Modela un problema de extracción de datos desde bases de datos no estructuradas. 3. Reconoce el concepto de datos estructurados y no estructurados. 4. Aplica las técnicas de extracción de datos estructurados, semiestructurados y no estructurados. 5. Aplica y evalúa técnicas de

Unidades	Contenidos	Resultados de aprendizaje
		optimización de consultas para extracción de datos. 6. Formula y evalúa el desempeño de las consultas para extracción de datos.
3. Extracción de Datos desde APIs y Fuentes Web.	3.1. Fundamentos de la extracción de datos desde APIs. 3.2. Técnicas de Web Scraping. 3.3. Consideraciones éticas y legales en la extracción de datos web.	1. Reconoce conceptos de extracción de datos mediante APIs 2. Comprende el funcionamiento de un API para integración de datos 3. Reconoce técnicas de web scraping para extracción automática de datos vía web 4. Aplica técnicas de web scraping para extracción automática de datos vía web 5. Comprende consideraciones éticas y legales en relación a la extracción de datos de fuentes de datos internas y externas.
4. Procesamiento y Limpieza de Datos.	4.1. Verificación de integridad de datos, identificación de datos duplicados e inconsistentes. 4.2. Técnicas de limpieza de datos. 4.3. Transformación de datos para análisis.	1. Reconoce y comprende técnicas de verificación de integridad de datos. 2. Modela un proceso de verificación de integridad de datos. 3. Reconoce técnicas de limpieza, verificación de integridad, consistencia y transformación de datos. 4. Aplica las técnicas de limpieza, verificación de integridad y consistencia de datos. 5. Aplica técnicas de transformación de datos con orientación al análisis. 6. Formula y evalúa el desempeño de procesos de limpieza, transformación y verificación de integridad y consistencia de datos.
5. Herramientas y Plataformas de Extracción de Datos.	5.1. Uso de herramientas como Python (pandas), R, Talend, y otras. 5.2. Implementación de procesos ETL en entornos de Big Data (Hadoop, Spark). 5.3. Evaluación de rendimiento y eficiencia de los procesos ETL.	1. Reconoce herramientas computacionales disponibles para extracción de datos. 2. Reconoce entornos de Big Data y comprende su funcionamiento. 3. Aplica herramientas de extracción de datos y entornos de Big Data para la implementación de procesos de ETL. 4. Formula y evalúa el desempeño en cuanto a rendimiento y eficiencia de un proceso de ETL.



V. ESTRATEGIAS DIDÁCTICAS

En el desarrollo del programa se aplicarán estrategias didácticas conducentes a la apropiación teórica y la ejecución práctica de procesos y procedimientos, a saber:

- **Debate:** exposición por parte del docente de los conceptos básicos por unidad, con materiales de lectura y ejemplos orientados a la enseñanza de las competencias específicas de la asignatura. El docente asume el rol de expositor y buscará generar el debate a través de preguntas sobre lo expuesto y desde la participación de los estudiantes.
- **Trabajos Colaborativos:** se realizarán trabajos prácticos por temáticas en equipos, se busca la cooperación y el trabajo grupal, incentivando la designación de roles y responsabilidades dentro de un equipo de trabajo, el compromiso grupal para el logro del objetivo del trabajo y la práctica de aplicación de la teoría a problemas reales para permitir validar la teoría con la práctica.
- **Estudios de casos:** se presentan varios estudios de casos de investigación que han aplicado diferentes técnicas estudiadas, permite identificar al estudiante un escenario real de aplicación de las técnicas aprendidas, comprender la manera de cómo encarar la resolución de un problema en un escenario real y aprender en base a la experiencia previa de trabajos de investigación publicados.

La elección particular de la estrategia didáctica aplicada será explícita en el Planeamiento de la Asignatura, de acuerdo con el perfil de los estudiantes, los recursos disponibles y el contexto educativo.

VI. ESTRATEGIAS EVALUATIVAS

Procesos de producción grupales e individuales, pruebas individuales orales y/o escritas durante el desarrollo de las unidades con diálogos e interpretaciones que los estudiantes realicen sobre los contenidos, debates, retroalimentación en casos necesarios y actividades que amplíen el conocimiento.

Con fines de calificación y promoción se aplicará el Reglamento Académico vigente en la institución que prevé valoraciones de proceso y final.

VII. MEDIOS AUXILIARES

Pizarras, marcadores, computadoras, proyector, parlantes para multimedia, plataforma virtual, sala de laboratorio equipada para prácticas con sistemas operativos Linux o Windows y acceso a Internet.

VIII. BIBLIOGRAFÍA

- Kimball, R., Caserta, J. (2004). The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data. Wiley.
- Densmore, J. (2021). Data Pipelines Pocket Reference: Moving and Processing Data for Analytics. O'Reilly Media
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann.
- Cunningham, J. P., & Ghahramani, Z. (2015). Linear Dimensionality Reduction: Survey, Insights, and Generalizations. Journal of Machine Learning Research, 16(1), 2859-2900.
- Reis, J. Housley, M. (2022). Fundamentals of Data Engineering: Plan and Build Robust Data Systems. O'Reilly Media
- Kleppmann, M. (2017). Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems. O'Reilly Media
- Crickard, P. (2020). Data Engineering with Python. Packt Publishing.

