Ministerio de Cultura y Educación
Universidad Nacional de San Luis
Facultad de Ciencias Físico Matemáticas y Naturales
Departamento: Informatica
Area: Area II: Sistemas de Computacion

(Programa del año 2024)
(Programa en trámite de aprobación)
(Presentado el 18/04/2024 11:21:42)

I - Oferta Académica

Materia	Carrera	Plan	Año	Período
LABORATORIOS DE DATOS	LICENCIATURA EN ANÁLISIS Y GES	OCS-1-27/22	2024	1° cuatrimestre

II - Equipo Docente

Docente	Función	Cargo	Dedicación
BARRIONUEVO, MERCEDES DEOLINDA	Prof. Responsable	P.Adj Exc	40 Hs
ARROYUELO BILLIARDI, JORGE A.	Prof. Co-Responsable	P.Adj Exc	40 Hs
TISSERA, PABLO CRISTIAN	Prof. Co-Responsable	P.Adj Exc	40 Hs

III - Características del Curso

Credito Horario Semanal				
Teórico/Práctico	**Teóricas**	**Prácticas de Aula**	**Práct. de lab/ camp/ Resid/ PIP, etc.**	**Total**
2 Hs	Hs	Hs	4 Hs	6 Hs

Tipificación	Periodo
B - Teoria con prácticas de aula y laboratorio	1º Cuatrimestre

Duración			
Desde	**Hasta**	**Cantidad de Semanas**	**Cantidad de Horas**
11/03/2024	21/06/2024	15	90

IV - Fundamentación

La gran cantidad de datos recopilados y almacenados por las distintas tecnologías actuales tienen la capacidad de generar beneficios transformadores para toda organizaciones, sin embargo, para ello debemos ser capaces de poder unificarlos en un formato específico, limpiarlos e interpretarlos.
Por lo tanto, resulta necesario abordar los aspectos de ETL (Extracción, Transformación y Carga), creación y manipulación de base de datos relacionales y visualización e interpretación de datos.
En esta asignatura se sientan las primeras bases que serán utilizadas por diversas asignaturas de años superiores.

V - Objetivos / Resultados de Aprendizaje

El objetivo general de la asignatura es suministrar al alumno conocimientos sólidos referidos al campo de los datos, examinando su amplio conjunto de principios fundamentales, tecnologías y aplicaciones para almacenar, transformar y visualizar datos para colaborar con los usuarios en la reducción de incertidumbre y toma de decisiones fundamentadas.
Se espera que al finalizar el curso el alumno sea capaz de:
* Manejar con idoneidad los conceptos que involucran el proceso de ETL, almacenamiento, visualización e interpretación de datos.
* Comprender el uso y el campo de aplicación de las distintas tecnologías asociadas.

* Desarrollar una actitud crítica frente al uso de las estructuras de manipulación de datos.

VI - Contenidos

Unidad I: Introducción a los datos.
Introducción. El dato como unidad de análisis. Tipos de datos. Características de los datos. Información. Conocimiento. Definición de Bases de Datos. El concepto de esquema e instancia. Sistemas de Gestión de Bases de Datos (SGBD).

Unidad II: Preprocesamiento y Análisis Exploratorio de Datos.
Limpieza de datos (data cleaning). Análisis exploratorio de datos. Análisis univariado y bivariado. Ruidos y anomalías (Outliers).
Normalización de datos. Discretización y Binarización de variables.
Análisis de las principales Herramientas ETL disponibles.

Unidad III: Almacenamiento en Base de Datos
Creación, modificación y almacenamiento de bases de datos. Tablas, índices, generación de scripts. Programar consultas, manipular datos y administrar índices.

Unidad IV: SQL (Structure Query Language)
Estudio de SQL. El lenguaje de consultas SQL. Conjunto de instrucciones básicas del SQL: create table, drop, update, insert, delete y select. Funciones agregadas. Consultas sumarias: group by y having.

Unidad V: Visualización de datos
Definición de Visualización de datos. El proceso de la Visualización. Estados y transformaciones de los datos. Variables Visuales. Interacciones. Dashboards.
Aplicaciones prácticas en el campo técnico y social con software estadístico para el análisis y la descripción de datos. Visualización, análisis e interpretación. Aplicación en el campo técnico y social.

VII - Plan de Trabajos Prácticos

Se propiciará el aprendizaje de los alumnos no sólo a través de clases teóricas con instancias de participación, sino también de clases de trabajos prácticos de laboratorio; haciendo hincapié en el Aprendizaje Basado en Proyectos (ABP).
Partiendo de un conjunto de datos a elección del alumno o proporcionado por la cátedra al cual se le aplicará cada una de las etapas del proceso de extracción de conocimiento para así obtener algún conocimiento buscado. Asimismo, se propiciará el trabajo grupal colaborativo y proactivo.

Trabajos Prácticos de máquina/laboratorio

Laboratorio 1: Utilización de herramientas que permitan el análisis exploratorio y pre-procesamiento de datos
Laboratorio 2: Instalación de una base de datos relacional. Utilización de lenguaje SQL para responder consultas.
Laboratorio 3: Utilización de herramientas de visualización para la muestra de resultados

Los laboratorios 1, 2 y 3 conformaran el trabajo integrador a presentar al finalizar la materia.

VIII - Regimen de Aprobación

El curso se implementa con condiciones específicas para estudiantes promocionales, regulares y libres (Ordenanza C.S. 13/03 y modificatoria Ordenanza C.S. 32/14).

Régimen para Regularizar el curso:
1. Participar al menos en un 60% en los foros obligatorios.
2. Aprobar los Trabajos Prácticos Evaluativos (TPE) con al menos el 60%.
3. Aprobar el trabajo final integrador con un 60%.
4. Aprobar el examen parcial con al menos el 60% o sus respectivas recuperaciones. Se darán 2 recuperaciones.
Si cualquier punto no fuese cumplimentado, implica que el alumno pase a condición de libre.

Régimen para promocionar el curso:
1. Participar al menos en un 80% en los foros obligatorios.
2. Aprobar los Trabajos Prácticos Evaluativos (TPE) con al menos el 70%.
3. Aprobar el trabajo final integrador con al menos el 70%.
4. Aprobar el examen parcial con al menos el 70% o sus respectivas recuperaciones. Se darán 2 recuperaciones.

4. La nota final se computará promediando las notas obtenidas en cada uno de los puntos mencionados previamente.

Trabajos Prácticos Evaluativos (TPE):
Los contenidos evaluados en estos TPE de conocimiento serán:
• TPE nº1: Unidades 1 y 2: Dato, Información y ETL
• TPE nº2: Unidades 3 y 4 Base de datos y lenguaje SQL
• TPE nº2: Unidades 5: Visualización de datos.
Los TPE programados serán virtuales e individuales y, cada uno de ellos, tendrá 2 (dos) recuperaciones.

-
Examen final para estudiantes regulares:
El examen final para estudiantes regulares consistirá en una evaluación escrita (virtual) teórico-práctica, con temas de todas las unidades previstas en el programa, y se aprueba con un porcentaje igual o superior al 60%.

Examen Libre:
No se admiten alumnos libres dado que los prácticos de máquina y laboratorios se desarrollan de manera incremental desde comienzo de cuatrimestre, por consiguiente no es posible en un examen poder evaluar correctamente este proceso.

IX - Bibliografía Básica

[1] Apuntes de cátedra.
[2] Data Science Concepts and Techniques with Applications. Usman Qamar, Muhammad Summair Raza. Springer 2020, ISBN: 978-981-15-6132-0.
[3] Texts in Computer Science, The Data Science Design Manual. Steven S. Skiena. Springer 2017. ISBN: 978-3-319-55443-3.
[4] Probabilidad y Estadística para Ingeniería y Ciencias. 9th Edición. Ronald E. Wallpole, Raymond H. Myers, Sharon L. Myers and Keying Ye. Pearson Education, 2012. ISBN: 978-607-32-1417-9.
[5] Hernández Orallo, Ramírez Quintana, Ferri Ramírez. Introducción a la Minería de Datos. Prentice Hall. 2004. ISBN 84-205-4091-9.
[6] Ian H. Witten, Eibe Frank, Mark A. Hall. Data Mining: Practical Machine Learning Tools and Techniques, (Fourth Edition). Morgan Kaufmann. 2017. ISBN 978-0-12-804291-5.
[7] Charu C. Aggarwal . Data Mining: The Textbook. Springer, 2015. ISBN 978-3319141428.
[8] Visualization of Time-Oriented Data. Springer Publishing Company,
[9] Incorporated, 1st edition, 2011.
[10] Tamara Munzner and Eamonn Maguire. Visualization analysis and design. A K Peters visualization series. CRC Press, Boca Raton, FL, 2015.

X - Bibliografia Complementaria

[1]

XI - Resumen de Objetivos

Manejar con idoneidad los conceptos que involucran el proceso de ETL, almacenamiento, visualización e interpretación de datos.

XII - Resumen del Programa

El dato como unidad de análisis.

Preprocesamiento y análisis exploratorio de datos
Almacenamiento en Base de datos
SQL (Structure Query Language)
Visualización de datos

XIII - Imprevistos

XIV - Otros

Mail del docente responsable
mdbarrio@email.unsl.edu.ar

ELEVACIÓN y APROBACIÓN DE ESTE PROGRAMA	
	Profesor Responsable
Firma:	
Aclaración:	
Fecha:	