



Formato para registro de Unidades de aprendizaje 2021

I.- Datos de identificación de la unidad de aprendizaje

Unidad académica:	Multisede (CIC, CIDETEC, ESCOM, ESFM, UPIITA)									
Programa académico:	Doctorado en Ciencia y Tecnología de Inteligencia Artificial y Ciencia de Datos									
	X	Doctorado		Orientación profesional						
		Maestría	X	Orientado a la investigación						
		Especialidad		Con la industria						
				Especialidad médica						
Nombre de unidad de aprendizaje:	Sesión de colegio donde se propuso:					Fecha de propuesta:				
	Acondicionamiento y visualización de datos									
	Clave de la unidad de aprendizaje:					Créditos:		5		REP 2017
	Semanas del semestre		18		Horas a la semana:		4		Horas totales: 72	
Tipo de unidad de aprendizaje:	Obligatoria:				Optativa:		X		Observaciones:	
	Semestre:		1 - 4							
	Teórica (%):		50		Práctica (%):		50		Teórico-prácticas (%):	
Área del conocimiento:	Ingeniería y Ciencias Fisicomatemáticas		X		Ciencias Sociales y Administrativas				Ciencias Médico Biológicas	
									Interdisciplinario	
Modalidad no escolarizada:	No escolarizada				Nombre de la Plataforma:					
	Mixta				Presencial (%):				En plataforma (%):	
Horas establecidas en el programa de estudios:	Presenciales (si procede) (horas x semana)								En plataforma (horas x semana):	



Formato para registro de Unidades de aprendizaje 2021

I. Aprendizajes que el estudiante deberá demostrar al finalizar

Conocimientos	Habilidades y destrezas	Actitudes y valores
• Diferenciar las técnicas para medir la distancia o similitud a utilizar dependiendo del tipo de datos para detectar la redundancia en los conjuntos de datos. • Usar técnicas que ayudan a reducir el número de dimensiones acorde al tipo de los datos a analizar y visualizar. • Saber las particularidades en cuanto a la utilidad requerida de los diferentes principales tipos de visualizaciones a emplear para analizar y entregar resultados que apoye la toma de decisiones. • Conocer las herramientas para realizar visualizaciones en turno, por sus ventajas y desventajas.	• Describir con facilidad las características importantes (tipos, volumen, complejidad entre otras) de los conjuntos de datos a analizar • Identificar las posibles visualizaciones a usar para desarrollar el análisis de datos requerido. • Ligar las técnicas para medir las similitudes o diferencias, de reducción de dimensiones y de las visualizaciones para analizar el/los conjunto(s) de datos. • identificar el conocimiento qué es relevante o interesante para el conjunto de datos con ayuda del dueño de los datos	• Innovación y creatividad para aprovechar la información que se genera día con día a través de diferentes medios • Respetar la información confidencial y el conocimiento derivado de los conjuntos de datos de los dueños de ella, sin importar si son, personas, empresas y competidores. • Utilizar únicamente información disponible y legal

Resolución que aborda la propuesta con su enfoque disciplinar

La visualización está transformando el análisis exploratorio de datos. Los analistas de datos se ocupan de conjuntos de datos más grandes que nunca, lo que dificulta el análisis. El análisis de datos visual, respaldado por crecientes capacidades y los métodos de la Ciencia de Datos o Minería de Datos de la computación mejora sus resultados. El análisis exploratorio de datos (EDA), un término acuñado por el renombrado estadístico John Tukey, es una técnica para comprender inicialmente y desarrollar una vista de un conjunto particular de datos antes de que comience la investigación profunda. La visualización con las técnicas estadísticas para describir las características de los datos ayuda al fin de generar hipótesis iniciales. En la era de los macrodatos, cuando los conjuntos de datos crecen rutinariamente hasta peta bytes, la visualización es más importante que nunca. La mayoría de los depósitos de información de hoy en día son simplemente demasiado grandes, complejos y diversos para explorarlos mediante un análisis numérico de memoria.

Sin importar el tipo de los datos a analizar, los analistas se enfrentan a desafíos adicionales. Estos conjuntos de datos pueden incluir cambios a lo largo del tiempo (datos espaciotemporales); la capacidad de caracterizar y obtener rápidamente una perspectiva de tales datos multidimensionales es un paso esencial, particularmente en situaciones críticas como la respuesta a desastres u otro fenómeno crítico a analizar.



Formato para registro de Unidades de aprendizaje 2021

Las capacidades de análisis visual intuitivas y de alto rendimiento serán vitales para el éxito de la visualización en medio de esta creciente ola de información. Los humanos están predispuestos a comprender información compleja en términos visuales. Los mapas son un ejemplo perfecto; las formas, los colores, las líneas, los puntos de referencia y la escala comparativa son formas en las que los humanos captan rápidamente el significado de los mapas.

La presentación visual es donde las hojas de cálculo se quedan cortas. Realizar un análisis utilizando software de hoja de cálculo implica filtros y fórmulas que consumen mucho tiempo. Las hojas de cálculo carecen de la velocidad, tanto en el cálculo como en la comunicación, para permitir que los analistas identifiquen rápidamente las relaciones y generen ideas.

La visualización a escala también permite a los analistas trabajar a la velocidad de su curiosidad natural. En la mayoría de los escenarios profesionales, los datos ya no se tratan de generar un informe; se trata de aprender. Los analistas necesitan herramientas que ayuden a determinar qué preguntas se deben hacer y, en función de las respuestas, descubrir cuáles deben ser las siguientes preguntas. Esta forma natural de consulta, posible gracias al análisis de datos de alto rendimiento, no solo unirá sistemas incompatibles, sino que también estimulará el conocimiento y la comprensión.

- **Jiawei Han and Micheline Kamber**, “Data Mining: Concepts and Techniques”, chapters two and three, Third edition The Morgan Kaufmann Series in Data Management Systems, Jim Gray, Series Editor, [Morgan Kaufmann Publishers](http://www.morgankaufmannpublishers.com), March 2011. ISBN 1-55860-901-6
- <https://www.kdnuggets.com/2021/08/visualization-transforming-exploratory-data-analysis.html>

II. Proximidad formativa

Áreas multi, inter y transdisciplinarias	Líneas de Generación y Aplicación de Conocimiento	Sectores sociales
• Lógica matemática • Probabilidad y estadísticas • Matemáticas para análisis • Modelado e inferencia • Informática • Algoritmos	• Aprendizaje automático • Redes neuronales y aprendizaje profundo • Reconocimiento de patrones • Generación de soluciones y decisiones inteligentes: sistemas híbridos, búsqueda y optimización	• Los sectores que se establecen en las LGAC, agua, cultura, educación, energía y cambio climático, salud, seguridad, soberanía alimentaria y vivienda,



Formato para registro de Unidades de aprendizaje 2021

	• Minería de datos, descubrimiento de conocimiento y analítica avanzada • Visión computacional • Robótica inteligente • Sistemas difusos y razonamiento bajo incertidumbre • Minería de texto y procesamiento de lenguaje natural	
Estrategia de asociación: Trabajar con datos del libro de texto, con datos de bases de datos de ejemplos que acompañan a los sistemas administradores de datos y con datos abiertos que se publican en las páginas de organizaciones del gobierno, empresas privadas y organizaciones no gubernamentales (ONGs), las cuales integran y sensibilizan al estudiante de maestría Inteligencia Artificial y Ciencia de Datos en los problemas relacionados tanto a la sociedad como de empresas, y así aplicar los conocimientos impartidos en la U. de A.		

III Metodología de enseñanza – aprendizaje

Descripción
Se dará un seguimiento personalizado y de trabajo colaborativo mediante las diversas actividades que se realizan durante la unidad de aprendizaje con el fin de que el estudiante: - El trabajo lo desarrolle en las 18 semanas, en las 72 horas de estudio a distancia y/o presencial - Las sesiones están previstas para el seguimiento del trabajo colaborativo de los estudiantes, así como para ofrecerles apoyo en todas las dudas que tengan. - En el aspecto teórico, se desarrollarán se discutirán los 12 capítulos del libro de texto [1] - En el aspecto práctico, se revisará la aplicación de las técnicas a los datos que permitan abordar los problemas que estén relacionados con los sectores sociales ya enumerados en el punto anterior. Las evaluaciones se realizarán con base en la rúbrica correspondiente.



Formato para registro de Unidades de aprendizaje 2021

La calificación mínima aprobatoria para la Unidad de aprendizaje será ocho (8). Siendo la escala de notas y valores numéricos la siguiente:

Menor a 8: No aprobado

8: Bueno

9: Muy bueno

10: Excelente

Evidencias como proceso de aprendizaje

- Lista de asistencia en las 18 semanas
- Lista de dudas registradas en las asesorías solicitadas
- Entrega y discusión de los ejercicios y proyectos de los 12 capítulos del libro de texto [1]

Entregas revisadas de las presentaciones y documentos relacionadas a las actividades descritas en el punto V. Secuencia programática y que cubren al contenido temático

Evidencias integradoras (resultados que contribuyen al curriculum)

Está centrado en el producto del portafolio de evidencias de aprendizaje (bitácora)

Ponderación

IV. Descripción de la participación esperada en el estudiante

Receptiva

Resolutiva

Autónoma

Estratégica



Formato para registro de Unidades de aprendizaje 2021

			Los proyectos están abiertos a lo que proponga el alumno.
--	--	--	---

Contenido temático

1. Recuperación de Información y Visualización (2 horas)
 - a) Visualización
 - b) Recuperación de la información
 - c) Perspectivas perceptuales y cognitivas de visualización
 - d) Visualización de la recuperación de información
2. Preliminares en la Recuperación de Información (10 horas)
 - a) Modelo de Espacio Vectorial
 - b) Métodos de Ponderación
 - c) Medidas de Similitud
 - d) Modelos de Información
 - e) Algoritmos de Agrupación
 - f) Evaluación de los resultados de recuperación
3. Modelos de visualización de múltiples puntos de referencia (6 horas)
 - a) Puntos de referencia múltiple
 - b) Modelo para puntos fijos múltiples de referencia
 - c) Modelos para múltiples puntos de referencia móviles
 - d) Modelo para la rotación automática de puntos de referencia



Formato para registro de Unidades de aprendizaje 2021

e) Implicación de Recuperación de Información

4. Característica Espacial Euclidiana Basada en Modelos de Visualización (6 horas)

- a) Espacio Euclidiano y sus características
- b) Introducción a los Modelos de Evaluación de Recuperación de Información
- c) Modelo de Visualización basado en la Distancia - Ángulo
- d) Modelo de Visualización basado en Ángulo - Ángulo
- e) Modelo de Visualización basado en la Distancia – Distancia

5. Mapas auto-organizados Kohonen - Una red neuronal artificial (6 horas)

- a) Introducción a las redes neuronales
- b) Mapas auto-organizados de Kohonen
- c) Implicación de la SOM en la Recuperación de Información

6. Red Asociativa Pathfinder (6 horas)

- a) Propiedades y descripciones de redes asociativas
- b) Implicaciones en la Recuperación de Información

7. Escalamiento Multidimensional (MDS) (6 horas)

- a) Descripciones de los métodos de análisis de MDS
- b) Implicaciones de las técnicas de MDS para la Recuperación de Información

8. Visualización de la Información en Internet (6 horas)

- a) Introducción
- b) Visualización de la información en Internet

9. La ambigüedad en Visualización de Información (6 horas)

- a) Ambigüedad y su implicación en la visualización de información
- b) Análisis de la ambigüedad en modelos de visualización de Recuperación de Información



Formato para registro de Unidades de aprendizaje 2021

10. La implicación de las metáforas en la Visualización de la Información (6 horas)

- a) Definición, elementos básicos y las características de una metáfora
- b) Fundamento cognitivo de las metáforas
- c) Los modelos mentales, las metáforas y la interacción persona-máquina
- d) Metáforas en la recuperación de la Visualización de Información
- e) Procedimientos y principios para aplicar la metáfora

11. Indicadores y criterios de evaluación para la Visualización en la Recuperación de Información (6 horas)

- a) Evaluación de la Recuperación de Información
- b) Benchmarks y normas de evaluación

12. Desafíos y tendencias (6 horas)

- a) Introducción
- b) Las comparaciones de los modelos de visualización
- c) Problemas y desafíos

V. Secuencia programática

No.	Tema	Objetivo de aprendizaje / competencia específica	Tiempo/Horas/Semanas	
1				
Actividad(es):			Tipo de interacción(es):	
			Referencias (s):	
Evidencia(s):				

Tipo de interacción: ID–Instrucción directa, TC–Trabajo colaborativo, AC–Análisis en campo, RP–Reflexión personal, PE–Presentación expositiva

Nota: *Replique esta sección las veces que sea necesario para cubrir toda la secuencia programática*

Indicar solo el número de las *Referencias* indizadas en la sección VII de este documento.



Formato para registro de Unidades de aprendizaje 2021

VI. Habilitadores tecnológicos

Disposiciones	Especificaciones / descripción de efectos
Conectividad	
Habilidades digitales	Internet para complementar la información del libro de texto, acceder a documentación de herramientas
Interoperabilidad	Tema de la unidad de aprendizaje, utilizado en la compartición de datos abiertos
Datos abiertos	Utilizados, para plantear diferentes análisis a las bases de datos abiertas seleccionadas
Big Data	Se usan los datos, pero no se generan más datos
Machine Learning	Uso de esta tecnología
Simulación	
Realidad aumentada	Posible uso de esta tecnología
Otro...	

Conferencias magistrales

1.
2.
3.

Notas complementarias

VII. Referencias

Documentales / electrónicas

4. Zhang, Jin; "Visualization for Information Retrieval", Springer, ISBN 978-3-540-75147-2, 2008
5. Chaomei Chen, "Visualization Information, beyond the horizon", Springer, ISBN-10: 184628340X, 2006
6. Ben Fry, "Visualizing Data", O'Reilly Media Inc, ISBN-10: 0-596-51455-7, ISBN-13: 978-0-596-51455-6, 2008
7. Julie Steele & Noah Iliinsky "Beautiful Visualization", O'Reilly Media, Inc. ISBN: 978-1-449-37987-2, 2010
8. Christopher D. Manning, "An Introduction to Information Retrieval", Cambridge University Press Cambridge, England, ISBN-10: 0521865719, 2008.



Formato para registro de Unidades de aprendizaje 2021

9. Bruce Croft, "Search Engines , Information Retrieval in Practice", Pearson, ISBN-10: 0136072240, 2009
10. Chaomei Chen, "Top 10 Unsolved Information Visualization Problems", Drexel University, 2005

VIII. Créditos y responsabilidades

Responsabilidad	Nombre completo	Clave de nombramiento /No. de empleado
Coordinador (Autor)	Gilberto Lorenzo Martínez Luna	16300-EH-22
Participante (Coautor)	Adolfo Guzmán Arenas	13702-EG-18/6/9500093
Asesor didáctico / Diseñador Instruccional	Elsa Rubio Espino	17133-EI-23
Tecnólogo educativo / Comunicólogo		
Corrector de estilo		
Programador multimedia / Diseñador gráfico		
Otro...		

VERIFICACIÓN GENERAL DE LA PLANEACIÓN DIDÁCTICA

Por la División de Operación y Promoción al Posgrado de la SIP

Nombre _____

FIRMA _____

REVISIÓN DE LA PLANEACIÓN DIDÁCTICA (VIABILIDAD)

Por la Subdirección de Diseño y Desarrollo de la DEV

Nombre _____

FIRMA _____



Instituto Politécnico Nacional

Secretaría Académica
Dirección de Educación Virtual

Secretaría de Investigación y Posgrado
Dirección de Posgrado

SIP-30

Formato para registro de Unidades de aprendizaje 2021

Por la Dirección de Posgrado Nombre _____ FIRMA _____ SELLO DE VALIDACIÓN	Por la Dirección para la Educación Virtual Nombre _____ FIRMA _____
--	--