



# UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN



# FIME

FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA

## UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

### Facultad de Ingeniería Mecánica y Eléctrica

#### PE Doctorado en Ingeniería de Sistemas

#### PROGRAMA ANALÍTICO

### I. Datos de Identificación de la Unidad de Aprendizaje:

1. **Clave y nombre de la Unidad de Aprendizaje:** PD123 Ciencia de los datos

2. **Frecuencia semanal:** horas de trabajo presencial 4

3. **Horas de trabajo extra aula por semana:** 2

4. **Modalidad:** ☒ Escolarizada ☐ No escolarizada ☐ Mixto

5. **Período académico:** ☒ Semestral ☐ Tetramestral ☐ Modular

6. **LGAC:** Sistemas estocásticos y simulación

7. **Ubicación semestral:** 1-8

8. **Área curricular:** formación, libre elección

9. **Créditos:** 4

10. **Requisito:** Ninguno

11. **Fecha de elaboración:** 20/01/2010

12. **Fecha de la última actualización:** 10/06/2020

13. **Responsable(s) del diseño:**

095808 Dr. Fernando López Irragarri

096633 Dra. Satu Elisa Schaeffer

## II. Presentación:

La *ciencia de datos* consiste en aplicar herramientas computacionales, métodos estadísticos y modelos matemáticos en general a conjuntos de datos multivariados, de múltiples tipos de entrada, posiblemente en diversos formatos, grandes cantidades, potencialmente conteniendo ruido, errores, omisiones y duplicados con la finalidad de poder obtener respuestas estadísticamente respaldadas a preguntas de interés.

## III. Propósito(s):

Que el estudiante se familiarice con los conceptos fundamentales de analítica de datos, conozca y sea capaz de usar las tecnologías más populares de analítica para resolver problemas reales.

## IV. Competencias del perfil de egreso:

### 14. Competencias del perfil de egreso

P1) Realizar investigación original y resolver problemas en el área de toma de decisiones en ambientes operativos que pueden ser dinámicos o inciertos para lograr una asignación más efectiva de recursos y decidir el curso de acción óptimo para lograr objetivos establecidos.

P2) Resolver problemas concretos en sistemas de la industria, la academia o el sector público en base a las herramientas de la toma de decisiones con bases científicas para lograr el mejor diseño, análisis, planeación o gestión de dichos sistemas.

### 15. Competencias generales a que se vincula la Unidad de Aprendizaje:

<i>Declaración de la competencia general vinculada a la unidad de aprendizaje</i>	<i>Evidencia</i>
C2) Utiliza los lenguajes lógico, formal, matemático, icónico, verbal y no verbal de acuerdo a su etapa de vida en el área de las ciencias para comprender, interpretar y expresar ideas, sentimientos, teorías y corrientes de pensamiento con un enfoque ecuménico.	Tareas
C3) Maneja las tecnologías de la información de acuerdo a los usos del campo de las ciencias y la comunicación como herramientas para el acceso a la información y su transformación en conocimiento, así como para el aprendizaje y trabajo colaborativo con técnicas de vanguardia que le permitan su participación constructiva en la sociedad.	Tareas,
C5) Emplea pensamiento lógico, crítico, creativo y propositivo, siguiendo los modelos de pensamiento científico para analizar fenómenos naturales y sociales que le permitan tomar decisiones pertinentes en su ámbito de influencia con responsabilidad social.	Tareas,
C11) Practica los valores promovidos por la UANL: verdad, equidad, honestidad, libertad, solidaridad, respeto a la vida y a los demás, respeto a la naturaleza, integridad, ética profesional, justicia y responsabilidad, en su ámbito personal y profesional para contribuir a construir una sociedad sostenible.	Tareas, proyecto
C14) Resuelve conflictos personales y sociales conforme a técnicas específicas en el ámbito académico y de su profesión para la adecuada toma de decisiones.	Tareas, proyecto



# UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN



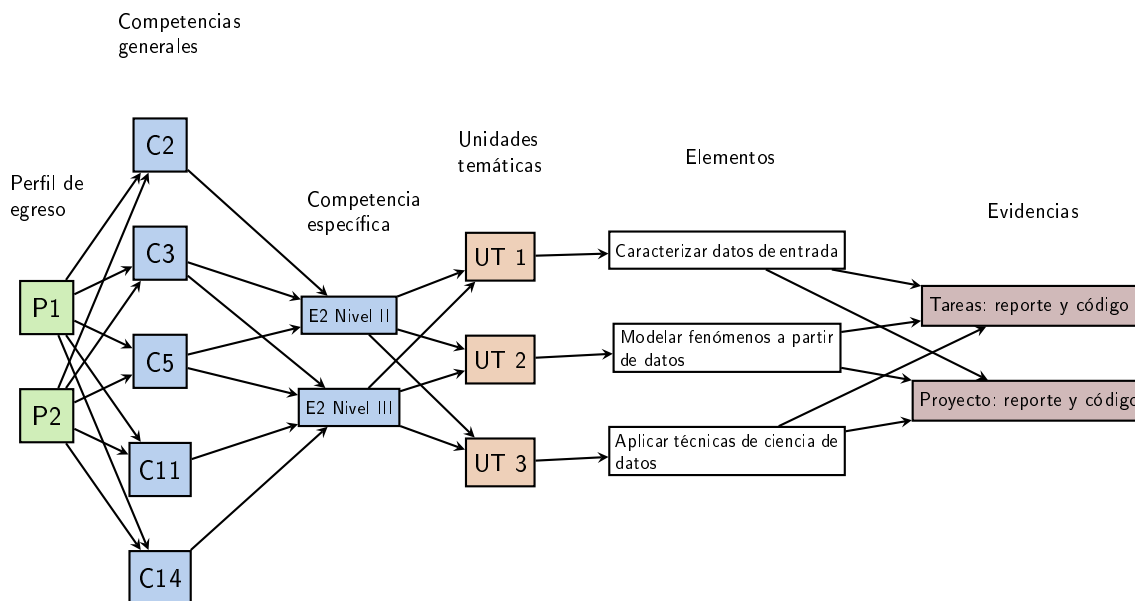
# FIME

FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA

## 16. Competencias específicas y nivel de dominio a que se vincula la unidad de aprendizaje:

Competencia Específica	Nivel I Inicial	Evidencia	Nivel II Básico	Evidencia	Nivel III Autónomo	Evidencia	Nivel IV Estratégico	Evidencia
E1) Realizar investigación original y resolver problemas en el área de toma de decisiones en ambientes operativos que pueden ser dinámicos o inciertos para lograr una asignación más efectiva de recursos y decidir el curso de acción óptimo para lograr objetivos establecidos.			Resuelve problemas de libro de texto en el área de toma de decisiones con bases científicas.	Tareas.	Encuentra soluciones para la consecución de objetivos establecidos para un problema dado, revisando literatura científica de frontera.	Tareas, proyecto.		

## V. Representación gráfica:



## VI. Estructuración en capítulos, etapas o fases de la unidad de aprendizaje:

### 17. Desarrollo de las fases de la Unidad de Aprendizaje:

Se busca desarrollar habilidades en la resolución en casos prácticos concretos. Se necesita contar con un buen entendimiento de varios los conceptos matemáticos, especialmente de matemáticas discretas y probabilidad, o en el caso contrario, estar preparado a estudiarlos según necesidad. También se necesita conocimiento de programación.

#### *Unidades temáticas*

**U1** Preprocesamiento y visualización (6 semanas)

**U2** Métodos estadísticos en la ciencia de datos (5 semanas)

**U3** Objetivos y técnicas para aplicaciones de la ciencia de datos (6 semanas)

La sesiones son de cuatro horas cada una y son veinte semanas en total. La primera semana es introductoria y las últimas dos semanas combinan elementos de las tres unidades temáticas en el contexto del proyecto integrador.

### *Temario semanal*

1. Introducción; selección de temas de proyecto
2. U1: Preparación de datos
3. U1: Lectura y manipulación de datos
4. U1: Estadística descriptiva básica
5. U1: Detección de anomalías
6. U1: Visualización estática de información
7. U1: Visualización dinámica de información
8. U2: Pruebas estadísticas
9. U2: Modelos lineales
10. U2: Regresión múltiple
11. U2: Análisis de varianza y de componentes principales
12. U2: Pronósticos
13. U3: Clasificación de datos
14. U3: Agrupamiento de datos
15. U3: Análisis de texto
16. U3: Análisis de imágenes
17. U3: Datos grandes
18. U3: Procesamiento en tiempo real
19. Presentaciones de proyectos
20. Revisión de portafolios de evidencia

### **Unidad temática 1: Preprocesamiento y visualización**

*Periodo:* 6 semanas

#### *Elementos de competencia:*

Evidencias de aprendizaje	Criterios de desempeño	Actividades de aprendizaje	Contenidos	Recursos
Seis (6) tareas semanales siendo cada una un reporte escrito y código de la implementación de un método de preprocesamiento y/o visualización de datos.	Calidad de la redacción científica del reporte; precisión del algoritmo propuesto; eficiencia de la implementación del método; cobertura de la experimentación.	Experimentación con ejemplos; lectura de material de apoyo; modificación de ejemplos; diseño y ejecución de experimentos; análisis y reporte de resultados obtenidos.	Métodos diversos de preprocesamiento y visualización de datos.	Material en la página web de la unidad y la literatura citada; lenguaje Python o similar; paquete $\text{\LaTeX}$ para redacción científica; repositorios de GitHub.

## Unidad temática 2: Métodos estadísticos en la ciencia de datos

Periodo: 5 semanas

Elementos de competencia:

Evidencias de aprendizaje	Criterios de desempeño	Actividades de aprendizaje	Contenidos	Recursos
Cinco (5) tareas semanales siendo cada una un reporte escrito y código de la implementación de un método estadístico de ciencia de datos.	Calidad de la redacción científica del reporte; precisión del método; eficiencia de la implementación del algoritmo; cobertura de la experimentación.	Experimentación con ejemplos; lectura de material de apoyo; modificación de ejemplos; diseño y ejecución de experimentos; análisis y reporte de resultados obtenidos.	Métodos estadísticos diversos de ciencia de datos.	Material en la página web de la unidad y la literatura citada; lenguaje Python o similar; paquete $\text{LATEX}$ para redacción científica; repositorios de GitHub.

## Unidad temática 3: Objetivos y técnicas para aplicaciones de la ciencia de datos

Periodo: 6 semanas

Elementos de competencia:

Evidencias de aprendizaje	Criterios de desempeño	Actividades de aprendizaje	Contenidos	Recursos
Seis (6) tareas semanales, siendo cada una un reporte escrito y código de la aplicación de un método de ciencia de datos.	Calidad de la redacción científica del reporte; precisión y eficiencia de la técnica aplicada; cobertura de la experimentación.	Experimentación con ejemplos; lectura de material de apoyo; modificación de ejemplos; diseño y ejecución de experimentos; análisis y reporte de resultados obtenidos.	Aplicaciones diversas de ciencia de datos.	Material en la página web de la unidad y la literatura citada; lenguaje Python o similar; paquete $\text{LATEX}$ para redacción científica; repositorios de GitHub.

## VII. Evaluación integral de procesos y productos:

Las tareas son individuales; se recomienda estudiar juntos y discutir las soluciones, pero no se tolera ningún tipo de plagio en absoluto, ni de otros estudiantes ni de la red ni de libros — toda referencia bibliográfica tiene que ser apropiadamente citada. La entrega se realiza por un repositorio público que debe reflejar todas las fases del trabajo.

No habrá examen. Son 17 tareas (A1–A17) que reportan avances semanales de aplicación de la lectura de la semana para el proyecto del estudiante, otorgando por máximo 5 puntos por tarea:

**NP** = tarea omitida

**5** = excede lo que se esperaba

**4** = cumple con lo que se esperaba

**3** = débil en alcance y/o calidad

**2** = débil en ambos alcance y calidad

**1** = sin contribuciones o méritos aunque fue entregada

**0** = completamente inadecuado en alcance y calidad

El proyecto final (A18) otorga un máximo de 15 puntos, evaluados en los siguientes rubros

1. Variedad de técnicas de empleadas
2. Cobertura y validez de la experimentación
3. Claridad y relevancia de los resultados
4. Calidad de visualización científica
5. Calidad de redacción científica

con la escala:

**3** = cumple con lo que se esperaba

**2** = débil en alcance y/o calidad

**1** = débil en ambos alcance y calidad

**0** = inadecuado en alcance y calidad

### Ponderación específica

Actividad																			Total
	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	
Ponderación	5 %	5 %	5 %	5 %	5 %	5 %	5 %	5 %	5 %	5 %	5 %	5 %	5 %	5 %	5 %	5 %	5 %	15 %	100 %



# UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN



# FIME

FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA

## VIII. Producto integrador de aprendizaje de la unidad:

### 18. Producto integrador de Aprendizaje:

Portafolio en un repositorio digital público que contiene los reportes escritos y los códigos de la implementación de todas las tareas y el proyecto integrador.

## IX. Fuentes de apoyo y consulta:

### 19. Fuentes de apoyo y consulta

#### Básicas

- D.T. LAROSE & C.D. LAROSE: *Data Science Using Python and R* Wiley, abril 2019, 256 páginas ISBN-13 978-1119526810
- H. WICKHAM & G. GROLEMUND: *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*, O'Reilly Media, enero 2017, 520 páginas, ISBN-13 978-1491910399
- M. TROVATI, et al., eds.: *Big-Data Analytics and Cloud Computing: Theory, Algorithms and Applications*. Springer, 2016.

**Complementarias** Artículos científicos especializados relacionados a los temas tratados, de preferencia publicados en revistas internacionales indizados recientes.