

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

COORDINACIÓN GENERAL DE FORMACIÓN PROFESIONAL

PROGRAMA DE UNIDAD DE APRENDIZAJE

I. DATOS DE IDENTIFICACIÓN

- 1. Unidad Académica:** Facultad de Ciencias Administrativas y Sociales, Ensenada; Facultad de Ciencias Administrativas, Mexicali; Facultad de Ciencias de la Ingeniería, Administrativas y Sociales, Tecate; y Facultad de Contaduría y Administración, Tijuana.
- 2. Programa Educativo:** Licenciado en Inteligencia de Negocios
- 3. Plan de Estudios:** 2021-2
- 4. Nombre de la Unidad de Aprendizaje:** Ciencia de Datos
- 5. Clave:** 39060
- 6. HC: 01 HT: 00 HL: 04 HPC: 00 HCL: 00 HE: 01 CR: 06**
- 7. Etapa de Formación a la que Pertenece:** Terminal
- 8. Carácter de la Unidad de Aprendizaje:** Obligatoria
- 9. Requisitos para Cursar la Unidad de Aprendizaje:** Ninguno



Equipo de diseño de PUA

Karina Caro Corrales
Roberto Carlos Valdés Hernández
Héctor Zatarain Aceves

Vo.Bo. de subdirectores de las Unidades Académicas

Adelaida Figueroa Villanueva
Angélica Reyes Mendoza
Esperanza Manrique Rojas
Jesús Antonio Padilla Sánchez

Fecha: 12 de marzo de 2021

II. PROPÓSITO DE LA UNIDAD DE APRENDIZAJE

La finalidad de esta asignatura es que el alumno integre los conocimientos, habilidades y herramientas para resolver problemas a través de la recolección, análisis y visualización de datos. Esto le permite identificar tendencias en los datos en diferentes campos de conocimiento y brindar información útil que coadyuve en el proceso de toma de decisiones estratégicas en las organizaciones. Se imparte en la etapa terminal con carácter obligatorio y pertenece al área de conocimiento de Ciencia de Datos.

III. COMPETENCIA GENERAL DE LA UNIDAD DE APRENDIZAJE

Analizar problemas basados en datos, utilizando herramientas y bibliotecas de la ciencia de datos, para recolectar, analizar y visualizar información relevante que permita apoyar en la toma de decisiones en las organizaciones, con honestidad, actitud analítica y propositiva.

IV. EVIDENCIA(S) DE APRENDIZAJE

Exposición de un proyecto de un estudio de caso en donde se haya utilizado la ciencia de datos, desde el planteamiento de algún problema a resolver, la recolección de datos, análisis y visualización de datos. Además, debe presentarse un reporte técnico que documente el proyecto.

V. DESARROLLO POR UNIDADES
UNIDAD I. Introducción a la ciencia de datos

Competencia:

Identificar los fundamentos de la ciencia de datos, a través de su aplicación y la distinción de sus diferentes roles, para comprender su utilidad en las organizaciones, con actitud analítica y crítica.

Contenido:

Duración: 2 horas

- 1.1 Qué es la ciencia de datos
- 1.2 Fundamentos de la ciencia de datos
- 1.3 Los múltiples caminos hacia la ciencia de datos
- 1.4 Roles en la ciencia de datos
 - 1.4.1. Científico de datos
 - 1.4.2. Ingeniero de datos
 - 1.4.3. Analista de datos

UNIDAD II. Recolección y tipos de datos

Competencia:

Clasificar los datos, mediante la recolección de los mismos, identificando sus atributos y visualización, para entender su comportamiento y prepararlos para su análisis, de forma inquisitiva y exacta.

Contenido:

Duración: 4 horas

- 2.1. Tipos y fuentes de datos
- 2.2. Recolección de datos
 - 2.2.1. Bases de datos públicas y privadas
 - 2.2.2. Screen scraping
- 2.2. Preparación de los datos
- 2.3. Representación y visualización de datos
- 2.3. Lenguajes para la analítica de datos
 - 2.3.1 Python, R y otros lenguajes

UNIDAD III. Analítica de datos

Competencia:

Examinar datos, mediante la aplicación de las diferentes técnicas de análisis y la selección de algoritmos pertinentes, para describir, explorar, inferir o predecir aspectos de interés que apoyen a la toma de decisiones en las organizaciones, con actitud reflexiva, sistemática y propositiva.

Contenido:

Duración: 6 horas

3.1. Tipos de análisis en la ciencia de datos

- 3.1.1. Descriptivo
- 3.1.2. Exploratorio
- 3.1.3. Inferencial
- 3.1.4 Predictivo
- 3.1.4. Causal
- 3.1.5. Mecanicista
- 3.1.6. Otros

3.2. Algoritmos de ciencias de datos

- 3.2.1 Algoritmos de clasificación
- 3.2.2. Algoritmos de clustering
- 3.2.3. Algoritmos de regresión
- 3.2.3. Algoritmos de aprendizaje de máquina
- 3.2.4. Otros

3.3. Herramientas, bibliotecas, servicios y APIs para la analítica de datos

- 3.3.1 IBM Watson Studio
- 3.3.2. R Studio
- 3.3.3. Tensor Flow con Python
- 3.3.4. Orange
- 3.3.5. Otras

UNIDAD IV. Visualización de datos

Competencia:

Aplicar técnicas de visualización en datos, considerando las diferentes categorías existentes y el uso de las diversas herramientas y bibliotecas de visualización, para comprender los resultados de la fase analítica previa, con creatividad y lógica.

Contenido:

Duración: 4 horas

- 4.1. Principios de la visualización de datos
 - 4.1.1. Cómo comunicar los hallazgos basados en datos
- 4.2. Categorías de visualización de datos
 - 4.2.1. Temporales
 - 4.2.2. Jerárquicas
 - 4.2.3. Redes
 - 4.2.4. Multidimensional
 - 4.2.5. Geoespacial
 - 4.2.6. Otras
- 4.3. Herramientas y bibliotecas de visualización
 - 4.3.1. Plotly
 - 4.3.2. ggplot2
 - 4.3.3. Tableau
 - 4.3.4. Microsoft Power BI
 - 4.3.5. Infogram
 - 4.3.6. D3
 - 4.3.7. Otras

VI. ESTRUCTURA DE LAS PRÁCTICAS DE LABORATORIO

| No. | Nombre de la Práctica | Procedimiento | Recursos de Apoyo | Duración |
|------------------|-----------------------|---|--|----------|
| UNIDAD II | | | | |
| 1 | Recolección de datos | <ol style="list-style-type: none"> 1. Atiende las orientaciones del profesor para indagar en las bases de datos públicas y privadas disponibles. 2. Recolecta datos de alguna base de datos pública o privada. 3. Examina los datos recolectados. 4. Clasifica y describe los datos recolectados. 5. Representa los datos recolectados mediante algún tipo de visualización. 6. Elabora un reporte con las actividades realizadas incluyendo clasificación, descripción y visualización de los datos. 7. Entrega el reporte al docente para su evaluación y retroalimentación. | <ul style="list-style-type: none"> ● Computadora con acceso a internet ● Acceso a bases de datos públicas y privadas ● Entorno de desarrollo/IDE ● Procesador de texto | 6 horas |
| 2 | Screen scraping | <ol style="list-style-type: none"> 1. Atiende las orientaciones del profesor para realizar el screen scraping de un sitio o aplicación web. 2. Selecciona la página o aplicación web. 3. Realiza el screen scraping utilizando un entorno de | <ul style="list-style-type: none"> ● Computadora con acceso a internet ● Entorno de desarrollo/IDE ● Procesador de texto ● Página o aplicación web | 6 horas |

| | | | | |
|---|---|--|---|---------|
| | | <p>desarrollo de programación (R, python, etc.).</p> <ol style="list-style-type: none"> Analiza la información recolectada mediante el screen scraping. Elabora un reporte con las actividades realizadas. Entrega el reporte al docente para su evaluación y retroalimentación. | | |
| 3 | Interfaces de programación de aplicaciones (APIs) | <ol style="list-style-type: none"> Atiende las orientaciones del profesor para utilizar una API para extraer datos de alguna aplicación. Selecciona la API a utilizar para realizar la extracción de datos (ej. Twitter API, Google Maps API, Facebook API, etc.). Utiliza la API seleccionada mediante un entorno de desarrollo para obtener un conjunto de datos de interés. Analiza y representa los datos recolectados mediante la API. Elabora un reporte con las actividades realizadas. Entrega el reporte al docente para su evaluación y retroalimentación. | <ul style="list-style-type: none"> Computadora con acceso a internet Entorno de desarrollo/IDE Librerías instaladas para trabajar con la API seleccionada Procesador de texto | 6 horas |
| 4 | Preparación y limpieza de datos | <ol style="list-style-type: none"> Atiende las orientaciones del profesor para realizar la preparación, limpieza y preprocesamiento de un conjunto de datos. Carga un conjunto de datos. Examina los datos. | <ul style="list-style-type: none"> Computadora con acceso a internet Entorno de desarrollo/IDE Conjunto de datos Procesador de texto | 6 horas |

| | | | | |
|-------------------|-------------------------------|--|--|---------|
| | | <ol style="list-style-type: none"> 4. Realiza la preparación, limpieza y preprocesamiento de los datos. 5. Representa el resultado de los datos preprocesados. 6. Elabora un reporte con las actividades realizadas. 7. Entrega el reporte al docente para su evaluación y retroalimentación. | | |
| UNIDAD III | | | | |
| 5 | Análisis descriptivo de datos | <ol style="list-style-type: none"> 1. Atiende las orientaciones del profesor para realizar un análisis descriptivo a un conjunto de datos. 2. Carga un conjunto de datos. 3. Examina los datos. 4. Realiza el análisis descriptivo al conjunto de datos utilizando un entorno de desarrollo de programación y diferentes librerías para realizar el análisis descriptivo. 5. Representa el resultado del análisis descriptivo. 6. Elabora un reporte con las actividades realizadas. 7. Entrega el reporte al docente para su evaluación y retroalimentación. | <ul style="list-style-type: none"> ● Computadora con acceso a internet ● Entorno de desarrollo/IDE ● Librerías instaladas para realizar el análisis ● Conjunto de datos ● Procesador de texto | 4 horas |
| 6 | Análisis inferencial de datos | <ol style="list-style-type: none"> 1. Atiende las orientaciones del profesor para realizar un análisis inferencial a un conjunto de datos. 2. Carga un conjunto de datos. 3. Examina los datos. 4. Establece las interrogantes a | <ul style="list-style-type: none"> ● Computadora con acceso a internet ● Entorno de desarrollo/IDE ● Librerías instaladas para realizar el análisis ● Conjunto de datos ● Procesador de texto | 4 horas |

| | | | | |
|---|------------------------------|--|--|---------|
| | | <p>resolver con el análisis.</p> <ol style="list-style-type: none"> 5. Realiza el análisis inferencial al conjunto de datos utilizando un entorno de desarrollo de programación y diferentes librerías para realizar el análisis descriptivo. 6. Representa el resultado del análisis descriptivo. 7. Elabora un reporte con las actividades realizadas. 8. Entrega el reporte al docente para su evaluación y retroalimentación. | | |
| 7 | Análisis predictivo de datos | <ol style="list-style-type: none"> 1. Atiende las orientaciones del profesor para realizar un análisis predictivo a un conjunto de datos. 2. Carga un conjunto de datos. 3. Examina los datos. 4. Establece las interrogantes a resolver con el análisis. 5. Realiza el análisis predictivo al conjunto de datos utilizando un entorno de desarrollo de programación y diferentes librerías para realizar el análisis descriptivo. 6. Representa el resultado del análisis descriptivo. 7. Elabora un reporte con las actividades realizadas. 8. Entrega el reporte al docente para su evaluación y retroalimentación. | <ul style="list-style-type: none"> ● Computadora con acceso a internet ● Entorno de desarrollo/IDE ● Librerías instaladas para realizar el análisis ● Conjunto de datos ● Procesador de texto | 4 horas |
| 8 | Algoritmos de Clasificación | <ol style="list-style-type: none"> 1. Atiende las orientaciones del profesor para aplicar un algoritmo de clasificación a un | <ul style="list-style-type: none"> ● Computadora con acceso a internet ● Entorno de desarrollo/IDE | 3 horas |

| | | | | |
|----|--------------------------------------|---|--|---------|
| | | <p>conjunto de datos.</p> <ol style="list-style-type: none"> 2. Carga un conjunto de datos. 3. Examina los datos. 4. Aplica un algoritmo de clasificación al conjunto de datos. 5. Analiza los resultados. 6. Elabora un reporte con las actividades realizadas. 7. Entrega el reporte al docente para su evaluación y retroalimentación. | <ul style="list-style-type: none"> ● Conjunto de datos ● Procesador de texto | |
| 9 | Algoritmos de clustering | <ol style="list-style-type: none"> 1. Atiende las orientaciones del profesor para aplicar un algoritmo de clustering a un conjunto de datos. 2. Carga un conjunto de datos. 3. Examina los datos. 4. Aplica un algoritmo de clustering al conjunto de datos. 5. Analiza los resultados. 6. Elabora un reporte con las actividades realizadas. 7. Entrega el reporte al docente para su evaluación y retroalimentación. | <ul style="list-style-type: none"> ● Computadora con acceso a internet ● Entorno de desarrollo/IDE ● Conjunto de datos ● Procesador de texto | 3 horas |
| 10 | Algoritmos de aprendizaje de máquina | <ol style="list-style-type: none"> 1. Atiende las orientaciones del profesor para aplicar un algoritmo de aprendizaje de máquina a un conjunto de datos. 2. Carga un conjunto de datos. 3. Examina los datos. 4. Aplica un algoritmo de aprendizaje de máquina al conjunto de datos. 5. Interpreta y analiza los | <ul style="list-style-type: none"> ● Computadora con acceso a internet ● Entorno de desarrollo/IDE ● Conjunto de datos ● Procesador de texto | 3 horas |

| | | | | |
|------------------|--|---|---|---------|
| | | <p>resultados.</p> <p>6. Elabora un reporte con las actividades realizadas.</p> <p>7. Entrega el reporte al docente para su evaluación y retroalimentación.</p> | | |
| 11 | Herramientas, Bibliotecas, servicios y APIs para la analítica de datos | <ol style="list-style-type: none"> 1. Atiende las orientaciones del profesor para utilizar una herramienta para la analítica de datos. 2. El docente proporciona un caso práctico 3. Aplica las funciones de analítica de datos de la herramienta (IBM Watson Studio, R Studio, Tensor Flow, Orange, etc.). 4. Interpreta y analiza los resultados más relevantes. 5. Elabora un reporte con las actividades realizadas. 6. Entrega el reporte al docente para su evaluación y retroalimentación. | <ul style="list-style-type: none"> • Computadora con acceso a internet • Herramienta para la analítica de datos (IBM Watson Studio, R Studio, Tensor Flow, Orange, etc.). • Conjunto de datos • Procesador de texto | 3 horas |
| UNIDAD IV | | | | |
| 12 | Categorías de visualización de datos | <ol style="list-style-type: none"> 1. Atiende las orientaciones del profesor para crear diferentes tipos de categorías de visualización de datos. 2. El docente proporciona un caso práctico 3. Carga el conjunto de datos. 4. Realiza los diferentes tipos de categorías de visualización de datos (Temporales, Jerárquicas, Redes, multidimensionales, | <ul style="list-style-type: none"> • Computadora con acceso a internet • Entorno de desarrollo/IDE • Conjunto de datos • Procesador de texto | 8 horas |

| | | | | |
|----|---|---|---|---------|
| | | <p>Geoespaciales, etc.).</p> <ol style="list-style-type: none"> 5. Interpreta y analiza los resultados más relevantes. 6. Elabora un reporte con las actividades realizadas. 7. Entrega el reporte al docente para su evaluación y retroalimentación. | | |
| 13 | Herramientas y bibliotecas de visualización | <ol style="list-style-type: none"> 1. Atiende las orientaciones del profesor para utilizar una herramienta para la visualización de datos. 2. El docente proporciona un caso práctico 3. Aplica las funciones de visualización de datos de la herramienta (Plotly, ggplot2, Tableau, Power BI, D3, etc.). 4. Interpreta y analiza los resultados más relevantes. 5. Elabora un reporte con las actividades realizadas. 6. Entrega el reporte al docente para su evaluación y retroalimentación. | <ul style="list-style-type: none"> ● Computadora con acceso a internet ● Herramienta para la visualización de datos (Plotly, ggplot2, Tableau, Power BI, D3, etc.). ● Conjunto de datos ● Procesador de texto | 8 horas |

VII. MÉTODO DE TRABAJO

Encuadre: El primer día de clase el docente debe establecer la forma de trabajo, criterios de evaluación, calidad de los trabajos académicos, derechos y obligaciones docente-alumno.

Estrategia de enseñanza (docente):

- Presenta información sobre los conceptos básicos de la ciencia de datos
- Presenta ejercicios prácticos relacionados con las temáticas
- Proporciona información para las prácticas de laboratorio
- Resuelve y ejemplifica con casos prácticos
- Dirige, supervisa y retroalimenta las prácticas de laboratorio
- Propicia la participación activa de los estudiantes
- Revisa y evalúa reportes de prácticas y actividades
- Diseña y aplica evaluaciones
- Muestra el uso de las herramientas de ciencia de datos
- Método de proyectos

Estrategia de aprendizaje (alumno):

- Indaga y analiza información sobre conceptos básicos de la ciencia de datos
- Realiza las prácticas de laboratorio
- Participa activamente en clase
- Trabaja de manera individual y en equipo
- Elabora y entrega actividades y reportes de prácticas en tiempo y forma
- Utiliza herramientas de ciencia de datos
- Atiende puntualmente las indicaciones del docente
- Elabora proyecto final

VIII. CRITERIOS DE EVALUACIÓN

La evaluación será llevada a cabo de forma permanente durante el desarrollo de la unidad de aprendizaje de la siguiente manera:

Criterios de acreditación

- Para tener derecho a examen ordinario y extraordinario, el estudiante debe cumplir con los porcentajes de asistencia que establece el Estatuto Escolar vigente.
- Calificación en escala del 0 al 100, con un mínimo aprobatorio de 60.

Criterios de evaluación

| | |
|---------------------------------|-------------|
| - Evaluaciones parciales..... | 20% |
| - Prácticas de laboratorio..... | 30% |
| - Tareas..... | 10% |
| - Proyecto..... | 40% |
| Total..... | 100% |

IX. REFERENCIAS

Básicas

- Adhikari, A., & DeNero, J. (2017). Computational and Inferential Thinking: The Foundations of Data Science. Recuperado de <https://www.inferentialthinking.com/chapters/intro>
- Albon, C. (2018). *Machine Learning with Python Cookbook: Practical Solutions from Preprocessing to Deep Learning*. Sebastopol, CA: O'Reilly Media.
- García, J., Molina, J. M., Berlanga, A., Miguel, P. Á., Bustamante, Á. L., Washington, P. R. (2018). *Ciencia de datos Técnicas Analíticas y Aprendizaje Estadístico*. España: Editorial Alfaomega.
- Géron, A. (2019). *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems* (2nd ed.). Sebastopol, CA: O'Reilly Media.
- Sánchez, R. (2017). *Ciencia de Datos con R. Importa Ordena Transforma Visualiza Comunica*. Editorial U.
- Skiena, S.S., (2017). *The Data Science Design Manual*. Springer. Recuperado de <https://link.springer.com/book/10.1007%2F978-3-319-55444-0>
- Zumel, N., Mount, J., & Porzak, J. (2020). *Practical data science with R*. (2nd ed.). Shelter Island, NY: Manning.

Complementarias

- Blum, A., Hopcroft, J., & Kannan, R. (2020). *Foundations of Data Science*. Cambridge University Press.
- Cady, F. (2017). *The data science handbook*. John Wiley & Sons.
- Estrellado, R. A., Freer, E. A., Mostipak, J., Rosenberg, J. M., & Velásquez, I. C. (2020). *Data science in education using R*. Routledge.
- Grus, J. (2019). *Data science from scratch: first principles with python*. O'Reilly Media.
- Kelleher, J. D., & Tierney, B. (2018). *Data science*. USA: MIT Press.
- Müller, A. C. & Guido, S. (2017). *Introduction to machine learning with Python: a guide for data scientists*. Japan: O'Reilly Media.

X. PERFIL DEL DOCENTE

El docente que imparta la unidad de aprendizaje de ciencia de datos deberá contar con título de licenciatura en inteligencia de negocios, informática, ingeniero en computación, sistemas computacionales, ciencias computacionales o área afín, maestría en área afín, con experiencia mínima de tres años en la docencia y profesional, y preferentemente que maneje las herramientas de software para la ciencia de datos. Ser proactivo, analítico, que fomente el trabajo en equipo y la investigación.