

1. Datos Generales de la asignatura

Nombre de la asignatura:	Ciencia de Datos
Clave de la asignatura:	SIC-2505
SATCA¹:	2-2-4
Carrera:	Ingeniería en Tecnologías de la Información y Comunicaciones

2. Presentación

Caracterización de la asignatura

Esta asignatura está diseñada para proporcionar a los estudiantes los conocimientos y habilidades necesarios para utilizar técnicas de programación avanzada en el contexto específico de la ciencia de datos.

La caracterización de esta asignatura se describe a través de los siguientes puntos clave:

- **Enfoque práctico y aplicado:** la asignatura se enfoca en desarrollar habilidades prácticas y aplicadas en el uso de técnicas de programación avanzada para resolver problemas reales en el ámbito de la ciencia de datos. Los estudiantes aprenderán a aplicar conceptos de programación avanzada en la manipulación, análisis, procesamiento y visualización de datos.
- **Interdisciplinariedad:** la asignatura integra conocimientos de diversas disciplinas, incluyendo la informática, la estadística y la ciencia de datos. Se explorarán métodos y herramientas de programación avanzada que son fundamentales para el análisis de datos y el desarrollo de modelos predictivos y descriptivos.
- **Profundización en conceptos de programación:** los estudiantes adquirirán una comprensión más profunda de conceptos de programación avanzada, como la programación orientada a objetos, la programación funcional, el manejo avanzado de datos y la implementación de algoritmos de machine learning. Se explorarán también técnicas de procesamiento de datos a gran escala y visualización avanzada.
- **Aplicabilidad en el mundo real:** la asignatura se centra en la aplicación práctica de los conocimientos adquiridos en contextos del mundo real. Se incluirán proyectos prácticos y casos de estudio que permitirán a los estudiantes enfrentarse a problemas reales de análisis de datos y encontrar soluciones utilizando herramientas y técnicas de programación avanzada.
- **Desarrollo de habilidades profesionales:** además de adquirir habilidades técnicas en programación avanzada para ciencia de datos, los estudiantes también desarrollarán habilidades profesionales importantes, como trabajo en equipo, resolución de problemas, comunicación efectiva y pensamiento crítico. Estas habilidades son fundamentales para el éxito en el campo de la ciencia de datos y la ingeniería.

Intención didáctica

La intención didáctica de esta materia es proporcionar a los estudiantes una formación sólida y práctica en el uso de técnicas de programación avanzada en el contexto específico de la ciencia de datos. Esta asignatura tiene como objetivo principal desarrollar en los estudiantes las habilidades necesarias para analizar, procesar y visualizar datos de manera efectiva utilizando herramientas y técnicas de programación avanzada, así como implementar y evaluar modelos de machine learning para la solución de problemas reales en diversos campos.

El tema 1, los estudiantes conocerán los fundamentos teóricos y prácticos de la ciencia de datos, así como su importancia y aplicación en diversos campos, desarrollando habilidades básicas para la manipulación de datos, análisis exploratorio y visualización.

En el tema 2, los estudiantes se familiarizarán con las principales herramientas de programación y desarrollo utilizadas en ciencia de datos, aprendiendo a usarlas para la manipulación, análisis y visualización de datos. Se enfocarán en el uso práctico de estas herramientas, permitiendo un flujo eficiente de trabajo en proyectos de ciencia de datos.

En el tema 3, se utilizarán las principales herramientas de visualización de datos, comprendiendo la importancia de las visualizaciones como parte fundamental del análisis de datos. Aprenderán a seleccionar y crear gráficos efectivos que comuniquen de manera clara y precisa los hallazgos obtenidos a partir del análisis de datos.

El tema 4, se comprenderán los desafíos éticos relacionados con el uso de Big Data, incluyendo la privacidad de los datos, el sesgo en los algoritmos, la responsabilidad en el uso de la información y las implicaciones sociales del análisis masivo de datos. Desarrollarán una conciencia crítica sobre cómo abordar estos dilemas éticos en proyectos de ciencia de datos.

¹ Sistema de Asignación y Transferencia de Créditos Académicos

3. Participantes en el diseño y seguimiento curricular del programa

Lugar y fecha de elaboración o revisión	Participantes	Evento
Instituto Tecnológico Superior de Puruándiro, 6 de septiembre al 10 octubre 2024	Docente de la Academia Ingeniería en Tecnologías de la Información y Comunicaciones.	Propuesta inicial.
Instituto Tecnológico Superior de Puruándiro, 15 de noviembre al 22 noviembre 2024	Docentes de la Academia Ingeniería en Tecnologías de la Información y Comunicaciones.	Reunión en academia para la revisión de los temas de la asignatura.

4. Competencia(s) a desarrollar

Competencia(s) específica(s) de la asignatura
Adquiere habilidades avanzadas en el uso de bibliotecas y herramientas de programación específicas para la ciencia de datos, como librerías para manejo de datos tabulares, visualización y entrenamiento de algoritmos de aprendizaje automático básicos.

5. Competencias previas

- Fundamentos de programación orientada a objetos. - Operaciones matriciales básicas. - Compresión de descriptores estadísticos básicos. - Obtención e interpretación de medidas y gráficos descriptivos. - Estimación puntual y por intervalos. - Planteamiento y prueba de hipótesis estadísticas. - Análisis, interpretación e informe de los resultados obtenidos en un estudio de inferencia estadística. - Analiza los requerimientos, manipula bases de datos para brindar soluciones al tratamiento de información basándose en modelos y estándares; de igual modo, implementa bases de datos para apoyar la toma de decisiones considerando las reglas de negocio.

6. Temario

No.	Temas	Subtemas
1	Introducción a la ciencia de datos	1.1 Introducción a la obtención de datos. 1.1.1. Definición y conceptos básicos de obtención de datos. 1.1.2. Importancia de la obtención de datos en Ciencia de Datos. 1.2. Web Scraping y APIs. 1.2.1. Web Scraping: recopilación de datos de sitios web 1.3. Uso de APIs avanzadas para datos especializados.
2	Herramientas de programación, desarrollo y análisis para ciencia de datos.	2.1. Lenguajes de programación para ciencia de datos. 2.2. Entornos de desarrollo de software. 2.3. Intérpretes en línea para desarrollo en ciencia de datos. 2.4 Acceso a archivos de formato especializado para ciencia de datos. 2.5 Uso de bibliotecas especializadas para operaciones matemáticas y científicas.
3	Herramientas para la visualización de datos.	3.1 Organización de datos para su análisis. 3.2 Representación tabular de datos. 3.3. Herramientas. 3.3.1. Tableau. 3.3.2. PowerBi. 3.3.4. ChartBlocks. 3.3.5 Datawrapper. 3.3.6 Gephi.
4	Ética y Big Data	4.1. Ética y responsabilidad en ciencia de Datos. 4.1.1. Privacidad de datos y protección del usuario. 4.1.2. Responsabilidad legal y regulaciones en ciencia de datos.

		4.2. Conceptos básicos a tecnologías para el procesamiento con Big Data. 4.3. Frameworks de procesamiento a gran escala 4.4. Modelos de aprendizaje supervisado con Big Data.
--	--	--

7. Actividades de aprendizaje de los temas

1. Introducción a la ciencia de datos	
Competencias	Actividades de aprendizaje
Específica(s): Desarrollar la capacidad para comprender los fundamentos teóricos y prácticos de la ciencia de datos, así como su importancia y aplicación en diversos campos. El estudiante será capaz de identificar, recolectar y procesar datos, realizar análisis exploratorio y presentar conclusiones mediante técnicas básicas de visualización de datos, con el fin de generar información útil para la toma de decisiones basadas en evidencia. Genérica(s): ● Aplica los conocimientos en la práctica, identificando aquellos que incorporen el compromiso con la responsabilidad social. ● Habilidades de gestión de información (habilidad para buscar y analizar información proveniente de fuentes diversas) ● Capacidad de aprender ● Solución de problemas Transversal(es): ● Aplica los conocimientos en la práctica, identificando aquellos que incorporen el compromiso con la responsabilidad social. ● Usa comunicación oral y escrita atendiendo los principios de no discriminación, Inclusión y equidad social. ● Diseña e implementa soluciones a problemas propios de ámbito de su área de aplicación integrando aprendizajes, rasgos y capacidades de excelencia, vanguardia e innovación social que fortalezcan el desarrollo humano	● Exploración de Datos (Data Exploration) ● Web Scraping: Recopilación de Datos de Sitios Web: Introducir a los estudiantes en la técnica de web scraping para extraer datos de sitios web y comprender su legalidad y ética. ● Interactuar con APIs especializadas para obtener datos que no están fácilmente disponibles en sitios web públicos. ● Scraping y Análisis de Sentimientos en Redes Sociales: Aplicar técnicas de scraping y análisis de texto para obtener y analizar datos de redes sociales.

2. Herramientas de programación, desarrollo y análisis para ciencia de datos.

Competencias Específica(s):	Actividades de aprendizaje
Diseñar entornos de desarrollo a partir de un análisis para la selección adecuada de software Genérica(s): ● Aplica los conocimientos en la práctica, identificando aquellos que incorporen el compromiso con la responsabilidad social. ● Habilidades de gestión de información (habilidad para buscar y analizar información proveniente de fuentes diversas) ● Capacidad de aprender ● Solución de problemas Transversal(es): ● Aplica los conocimientos en la práctica, identificando aquellos que incorporen el compromiso con la responsabilidad social. ● Usa comunicación oral y escrita atendiendo los principios de no discriminación, Inclusión y equidad social. ● Diseña e implementa soluciones a problemas propios de ámbito de su área de aplicación integrando aprendizajes, rasgos y capacidades de excelencia, vanguardia e innovación social que fortalezcan el desarrollo humano	• Investiga y evalúa diferentes lenguajes de programación populares para ciencia de datos, como Python, R o Julia. • Redacta un breve informe comparativo que destaque las fortalezas y debilidades de cada lenguaje, así como su idoneidad para diferentes tipos de proyectos de análisis de datos. • Investiga diferentes IDEs disponibles para el desarrollo de proyectos de ciencia de datos y elige uno para explorar en profundidad. • Realiza una lista de características clave de la IDE seleccionada y crea un breve tutorial que muestre cómo utilizar algunas de sus funcionalidades básicas.

3. Herramientas para la visualización de datos.	
Competencias	Actividades de aprendizaje
Especifica(s): Utiliza las herramientas de visualización y sabe distinguir en robustez y opciones cual es la que más se adapta a las diferentes situaciones que se le pueden presentar en futuras ocasiones. Genérica(s): ● Aplica los conocimientos en la práctica, identificando aquellos que incorporen el compromiso con la responsabilidad social. ● Habilidades de gestión de información (habilidad para buscar y analizar información proveniente de fuentes diversas) ● Capacidad de aprender ● Solución de problemas Transversal(es): ● Aplica los conocimientos en la práctica, identificando aquellos que incorporen el compromiso con la responsabilidad social. ● Usa comunicación oral y escrita atendiendo los principios de no discriminación, Inclusión y equidad social. ● Diseña e implementa soluciones a problemas propios de ámbito de su área de aplicación integrando aprendizajes, rasgos y capacidades de excelencia, vanguardia e innovación social	• Llevar a cabo actividades intelectuales de inducción-deducción y análisis-síntesis, las cuales lo encaminan hacia la investigación, la aplicación de conocimientos y la solución de problemas. • Buscar, seleccionar y analizar información en distintas fuentes. • Participar en actividades grupales que propicien la comunicación, el intercambio argumentado de ideas, la reflexión, la integración y la colaboración. • Llevar a cabo actividades prácticas que promuevan el desarrollo de habilidades para la experimentación, tales como: observación, identificación, manejo y control de variables y datos relevantes, planteamiento de hipótesis, trabajo en equipo. • Desarrollar actividades de aprendizaje que propicien la aplicación de los conceptos, modelos y metodologías que se van aprendiendo en el desarrollo de la asignatura.

que fortalezcan el desarrollo humano	
4. Ética y Big Data	
Competencias	Actividades de aprendizaje
Específica(s): Domina los conceptos fundamentales del procesamiento con Big Data, incluyendo la infraestructura tecnológica, herramientas y técnicas básicas para el manejo y análisis de grandes volúmenes de datos. Genérica(s): ● Aplica los conocimientos en la práctica, identificando aquellos que incorporen el compromiso con la responsabilidad social. ● Habilidades de gestión de información (habilidad para buscar y analizar información proveniente de fuentes diversas) ● Capacidad de aprender ● Solución de problemas	• Lecturas y estudio dirigido: lectura de material teórico sobre los fundamentos del procesamiento con Big Data, incluyendo conceptos, arquitecturas y tecnologías clave. • Prácticas en laboratorio: Realización de ejercicios prácticos utilizando herramientas como Hadoop, Spark o sistemas de bases de datos distribuidas para familiarizarse con el manejo de grandes volúmenes de datos. • Análisis de casos: Estudio y análisis de casos de estudio reales donde se aplicaron técnicas de procesamiento con Big Data para resolver problemas específicos en diferentes industrias. • Debates y discusiones: Participación en debates y discusiones en clase sobre temas relevantes relacionados con el

Transversal(es): ● Aplica los conocimientos en la práctica, identificando aquellos que incorporen el compromiso con la responsabilidad social. ● Usa comunicación oral y escrita atendiendo los principios de no discriminación, Inclusión y equidad social. ● Diseña e implementa soluciones a problemas propios de ámbito de su área de aplicación integrando aprendizajes, rasgos y capacidades de excelencia, vanguardia e innovación social que fortalezcan el desarrollo humano	procesamiento con Big Data, como desafíos éticos, impacto en la sociedad y tendencias futuras. ● Proyecto de investigación: Desarrollo de un proyecto de investigación o aplicación práctica que involucre el uso de herramientas y técnicas de procesamiento con Big Data, desde la recolección y preparación de datos hasta el análisis y la presentación de resultados.
--	---

--	--

8. Práctica(s)

Práctica de Obtención de Datos

Objetivo: Enseñar a los estudiantes cómo obtener y recolectar datos de diversas fuentes.

Descripción:

- Descargar datasets de repositorios públicos como Kaggle, UCI Machine Learning Repository o datos gubernamentales (open data).
- Obtener datos mediante APIs públicas (por ejemplo, API de Twitter, OpenWeather, etc.).
- Realizar scraping de páginas web usando herramientas como BeautifulSoup en Python.

Práctica de Limpieza y Preprocesamiento de Datos

- **Objetivo:** Aplicar técnicas para limpiar y preparar datos antes del análisis.
- **Descripción:**
 - Detección y tratamiento de valores nulos o faltantes.
 - Detección y eliminación de duplicados.
 - Transformación de datos categóricos en variables numéricas.
 - Normalización o escalado de datos.
 - Generación de nuevas variables a partir de las existentes (feature engineering).
- **Herramientas:** Python (pandas), R.

Práctica de Visualización de Datos

- **Objetivo:** Aprender a crear gráficos efectivos para visualizar patrones y comunicar resultados de manera clara.
- **Descripción:**
 - Creación de gráficos de barras, líneas, diagramas de dispersión, histogramas y mapas de calor.
 - Comparación de la efectividad de distintos tipos de gráficos.
 - Uso de bibliotecas de visualización interactivas (Plotly o Tableau) para crear dashboards dinámicos.
 - Presentación de datos en un informe visual.
- **Herramientas:** Python (matplotlib, seaborn, plotly), Tableau, Power BI.
- **Resultado esperado:** Un conjunto de gráficos y visualizaciones que representen de manera clara los hallazgos clave de un análisis.

9. Proyecto de asignatura

Objetivo del Proyecto:

El proyecto final de la asignatura de Ciencia de Datos tiene como propósito integrar los conocimientos adquiridos durante el curso, aplicando las diferentes fases del ciclo de vida de un proyecto de ciencia de datos a un problema real. Los estudiantes desarrollarán competencias en la obtención, análisis, modelado y visualización de datos, para finalmente comunicar resultados que permitan tomar decisiones basadas en datos.

Descripción General del Proyecto:

Los estudiantes deben trabajar en un problema práctico de ciencia de datos, eligiendo un conjunto de datos adecuado, y aplicar técnicas de limpieza, exploración, análisis predictivo, visualización y comunicación de resultados. Este proyecto puede realizarse de manera individual o en grupos (según el tamaño de la clase) y se espera que los estudiantes presenten tanto un informe escrito como una presentación final.

10. Evaluación por competencias

- Análisis de casos.
- Entrevistas a expertos
- Solución de problemas realizados en forma individual o en equipo.
- Discusiones y debates en equipos.
- Paneles de presentaciones de temas.
- Reportes de proyectos, investigaciones, trabajos, etc.
- Simulaciones y/o demostraciones.
- Esquemas gráficos (mapas conceptuales, mapas mentales, mapas

11. Fuentes de información

1. McKinney, W. (2022). Python for data analysis. " O'Reilly Media, Inc."
2. Hadley, W., & Garrett, G. (2016). R for data science: import, tidy, transform, visualize, and model data. O'Reilly Media, Inc.
3. Müller, A. C., & Guido, S. (2016). Introduction to machine learning with Python: a guide for data scientists. " O'Reilly Media, Inc."
4. Julie Steele, N. I. (2010). Beautiful Visualization: Looking at Data through the Eyes of Experts (Theory in Practice). EUA: O'Reilly Media; 1st edition.
5. Knafllic, C. N. (2015). Storytelling with Data: A Data Visualization Guide for Business Professionals. USA: Wiley; 1st edition.
6. White, T. (2012). Hadoop: The definitive guide. " O'Reilly Media, Inc."
7. Acharya, S., & Chellappan, S. (2015). Big Data and Analytics. Wiley
8. Kakarla, R., Krishnan, S., & Alla, S. (2021). Applied Data Science Using PySpark: learn the End-to-End Predictive Model-Building Cycle. Apress.