



## 1. Datos Generales de la asignatura

<b>Nombre de la asignatura:</b>	Tópicos selectos para ciencia de datos
<b>Clave de la asignatura:</b>	CDA-2423
<b>SATCA<sup>1</sup>:</b>	0-4-4
<b>Carrera:</b>	Ingeniería en Ciencia de Datos

## 2. Presentación

<b>Caracterización de la asignatura</b>
<p>Esta asignatura considera la interrelación de las materias de la carrera de Ciencia de Datos con el desarrollo social, económico y tecnológico del entorno laboral en que se encuentra inmerso el Instituto Tecnológico involucrado, dentro de un marco integral de conocimiento.</p> <p>En la actualidad, la necesidad de innovar en el área de ciencia de datos surge de la creciente complejidad y volumen de los datos generados en diversos ámbitos. La era digital ha llevado a un aumento exponencial en la cantidad y variedad de datos disponibles, desde datos estructurados en bases de datos hasta datos no estructurados como texto, imágenes y videos. Esta abundancia de datos plantea desafíos significativos en términos de almacenamiento, procesamiento y análisis.</p> <p>Además, la rápida evolución de la tecnología y las demandas cambiantes del mercado requieren soluciones innovadoras que aprovechen al máximo los datos disponibles para tomar decisiones informadas, descubrir conocimientos ocultos y prever tendencias futuras. La innovación en ciencia de datos es esencial para desarrollar nuevas técnicas, algoritmos y herramientas que permitan a las organizaciones extraer información valiosa de los datos y convertirla en ventajas competitivas tangibles. Esto incluye avances en áreas como el aprendizaje automático, el procesamiento de lenguaje natural, la inteligencia artificial y la analítica avanzada, entre otros, que pueden ayudar a abordar desafíos complejos y aprovechar oportunidades emergentes en un entorno empresarial en constante cambio.</p>
<b>Intención didáctica</b>
<p>La intención didáctica de una materia de Tópicos Selectos de Ciencia de Datos es proporcionar a los estudiantes una comprensión profunda y actualizada de los avances más recientes en el campo de la ciencia de datos, así como su relevancia y aplicación en la sociedad y la economía.</p> <p>En la primera unidad, que trata sobre ética y responsabilidad en ciencia de datos, se enfocará en presentar a los estudiantes dilemas éticos y responsabilidades asociadas con el uso de datos, algoritmos y modelos en la toma de decisiones, considerando implicaciones sociales, económicas y legales.</p>

<sup>1</sup> Sistema de Asignación y Transferencia de Créditos Académicos



La segunda unidad, Aplicaciones avanzadas de Aprendizaje Automático Profundiza en técnicas avanzadas de aprendizaje automático, como aprendizaje profundo, aprendizaje por refuerzo y aprendizaje semi-supervisado, aplicadas a problemas específicos del mundo real, considerando las posibles implementaciones en hardware y software específico del sistema

La tercera unidad examina las últimas tecnologías y tendencias en el campo de la ciencia de datos, los cuales pueden incluir, pero no limitados a computación cuántica, federated learning y sistemas de recomendación avanzados.

Por último, la cuarta unidad, Gestión y optimización de proyectos de Ciencia de Datos, aborda la gestión eficaz de proyectos de ciencia de datos, desde la definición del problema hasta la implementación de soluciones, considerando aspectos como planificación, coordinación de equipos, evaluación de riesgos y comunicación de resultados.

La intención didáctica de esta materia es proporcionar a los estudiantes una visión holística y actualizada de la ciencia de datos, permitiéndoles explorar los avances más recientes, las tendencias tecnológicas emergentes y el impacto social y económico de esta disciplina en la sociedad contemporánea. Se espera que al final del curso, los estudiantes estén equipados con el conocimiento y la perspectiva necesarios para enfrentar los desafíos y aprovechar las oportunidades en un campo en constante evolución.

### 3. Participantes en el diseño y seguimiento curricular del programa

Lugar y fecha de elaboración o revisión	Participantes	Observaciones
Instituto Tecnológico Superior de Alvarado del 21 al 23 agosto de 2023.	Representante del Instituto Tecnológico Superior de Alvarado.	Propuesta inicial.
Tecnológico Nacional de México 30 octubre 2023	Representante del Instituto Tecnológico de: Querétaro y del Instituto Tecnológico Superior de Alvarado.	Presentación de la propuesta de la carrera de Ingeniería en Ciencia de Datos.
Instituto Tecnológico de Querétaro Campus Norte del 19 al 22 de marzo 2024.	Representantes de los Institutos Tecnológicos de: Morelia, Puebla, Querétaro, Tehuacán. Instituto Tecnológico Superior de Alvarado. CENIDET. Representante de Ciencias Básica de los Institutos de: Celaya, Morelia y CIIDET.	Diseño y/o desarrollo curricular de la carrera de Ingeniería en Ciencia de Datos.



<p>Tecnológico Nacional de México del 22 al 24 de abril del 2024</p>	<p>Representante del Instituto Tecnológico de Querétaro e Instituto Tecnológico Superior de Alvarado.</p>	<p>Contraste y ajuste de las asignaturas de Ingeniería en Ciencia de Datos con respecto a las de Ing. en Inteligencia Artificial, Ing. en Desarrollo WEB e Ing. en Ciberseguridad</p>
<p>Tecnológico Nacional de México del 27 al 31 de mayo del 2024.</p>	<p>Representantes de los Institutos Tecnológicos de: Morelia, Querétaro. Instituto Tecnológico Superior de Alvarado. CENIDET.</p>	<p>Consolidación curricular de la carrera de Ingeniería Ciencia de Datos</p>

#### 4. Competencia(s) a desarrollar

<p><b>Competencia(s) específica(s) de la asignatura</b></p>
<ul style="list-style-type: none"> <li>● Identifica y aborda dilemas éticos en el uso de datos y algoritmos.</li> <li>● Aplica técnicas avanzadas de aprendizaje automático a problemas específicos.</li> <li>● Conoce tecnologías emergentes y su aplicación en proyectos de Ciencia de Datos.</li> <li>● Gestiona proyectos, desde la planificación hasta la implementación y comunicación de resultados.</li> <li>● Identifica y mitiga sesgos algorítmicos y problemas de equidad en modelos de datos.</li> <li>● Evalúa el impacto social, económico y legal de las decisiones basadas en datos.</li> <li>● Conoce las nuevas tecnologías y tendencias en constante evolución en el campo de la Ciencia de Datos.</li> </ul>

#### 5. Competencias previas

<ul style="list-style-type: none"> <li>● Diseña algoritmos de aprendizaje automático.</li> <li>● Construye aplicaciones utilizando visión artificial.</li> <li>● Desarrolla algoritmos de procesamiento de lenguaje natural.</li> <li>● Diseña bases de datos de acuerdo a necesidades de la aplicación</li> <li>● Conoce elementos fundamentales de computación en la nube.</li> <li>● Diseña procesos utilizando el paradigma de Internet de las cosas.</li> <li>● Analiza datos mediante el uso de estadística inferencial.</li> </ul>
---



## 6. Temario

No.	Temas	Subtemas
1	Ética y responsabilidad en ciencia de datos	<ul style="list-style-type: none"><li>1.1. Privacidad de datos y protección del usuario.</li><li>1.2. Sesgos algorítmicos y equidad en los modelos.</li><li>1.3. Transparencia y explicabilidad de los modelos.</li><li>1.4. Responsabilidad legal y regulaciones en ciencia de datos.</li><li>1.5. Impacto social de las decisiones basadas en datos.</li></ul>
2	Aplicaciones avanzadas de aprendizaje automático	<ul style="list-style-type: none"><li>2.1. Generación de imágenes mediante modelos generativos.</li><li>2.2. Redes neuronales recurrentes para procesamiento de lenguaje natural.</li><li>2.3. Algoritmos de aprendizaje por refuerzo para control inteligente.</li><li>2.4. Métodos de aprendizaje semi-supervisado para datos etiquetados escasos.</li><li>2.5. Técnicas de transferencia de aprendizaje y adaptación de dominio.</li></ul>
3	Tecnologías emergentes en ciencia de datos	<ul style="list-style-type: none"><li>3.1. Aplicaciones en seguridad de datos y trazabilidad.</li><li>3.2. Fundamentos de la computación cuántica en ciencia de datos.</li><li>3.3. Federated learning y colaboración segura en entornos distribuidos.</li><li>3.4. Sistemas de recomendación avanzados basados en modelos de grafos.</li><li>3.5. Desarrollos recientes en procesamiento de datos a gran escala.</li></ul>
4	Gestión y optimización de proyectos de ciencia de datos.	<ul style="list-style-type: none"><li>4.1. Definición de objetivos y alcance del proyecto.</li><li>4.2. Aplicación de metodologías ágiles y mejores prácticas en gestión de proyectos.</li><li>4.3. Evaluación y selección de herramientas y tecnologías adecuadas.</li><li>4.4. Estrategias de comunicación efectiva con partes interesadas.</li><li>4.5. Monitoreo y optimización continua del rendimiento del modelo en producción.</li></ul>



## 7. Actividades de aprendizaje de los temas

1. Ética y responsabilidad en ciencia de datos	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Comprende las diferentes fuentes y Identifica dilemas éticos en el uso de datos y algoritmos, considerando implicaciones sociales, económicas y legales.</p> <p><i>Genérica(s):</i></p> <ul style="list-style-type: none"> <li>● Capacidad de análisis y síntesis.</li> <li>● Capacidad de organizar y planificar.</li> <li>● Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>● Toma de decisiones.</li> <li>● Trabajo en equipo.</li> <li>● Habilidades de investigación.</li> <li>● Capacidad de generar nuevas ideas.</li> <li>● Liderazgo.</li> <li>● Habilidad para trabajar en forma autónoma.</li> <li>● Búsqueda del logro.</li> </ul>	<ul style="list-style-type: none"> <li>● Debate en clase sobre dilemas éticos en casos reales de uso de datos.</li> <li>● Análisis de casos de estudio para identificar sesgos algorítmicos y proponer soluciones.</li> <li>● Elaboración de un código de ética para proyectos de Ciencia de Datos.</li> </ul>
2. Aplicaciones avanzadas de aprendizaje automático	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i></p> <ul style="list-style-type: none"> <li>● Aplica técnicas avanzadas de Aprendizaje Automático a problemas específicos del mundo real, optimizando su capacidad de resolución de problemas.</li> </ul> <p><i>Genérica(s):</i></p> <ul style="list-style-type: none"> <li>● Capacidad de comunicación oral y escrita.</li> <li>● Habilidades del uso de las tecnologías de información.</li> <li>● Habilidades para buscar, procesar y analizar información procedente de fuentes diversas.</li> </ul>	<ul style="list-style-type: none"> <li>● Implementación de modelos de aprendizaje profundo en proyectos prácticos.</li> <li>● Desarrollo de un proyecto de investigación sobre técnicas de aprendizaje por refuerzo.</li> <li>● Experimentación con algoritmos de transferencia de aprendizaje en conjuntos de datos variados.</li> </ul>



<b>3. Tecnologías emergentes en ciencia de datos</b>	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Evalúa tecnologías emergentes en proyectos de Ciencia de Datos, se mantiene al tanto de las últimas tendencias en el campo.</p> <p><i>Genérica(s):</i></p> <ul style="list-style-type: none"> <li>● Capacidad de investigación.</li> <li>● Capacidad para identificar, plantear y resolver problemas.</li> <li>● Capacidad para tomar decisiones</li> <li>● Capacidad para aplicar los conocimientos en la práctica.</li> </ul>	<ul style="list-style-type: none"> <li>● Investigación guiada sobre aplicaciones de trazabilidad en la industria.</li> <li>● Diseño de un proyecto piloto utilizando computación cuántica para resolver un problema de optimización.</li> <li>● Desarrollo de un sistema de recomendación utilizando federated learning en un entorno distribuido.</li> </ul>
<b>4. Gestión y optimización de proyectos de ciencia de datos</b>	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Planifica e implementa estrategias para la comunicación de resultados, mejorando la capacidad de coordinación y ejecución.</p> <p><i>Genéricas:</i></p> <ul style="list-style-type: none"> <li>● Capacidad de investigación.</li> <li>● Capacidad para identificar, plantear y resolver problemas.</li> <li>● Capacidad para tomar decisiones.</li> <li>● Capacidad para aplicar los conocimientos en la práctica.</li> </ul>	<ul style="list-style-type: none"> <li>● Simulación de la gestión de un proyecto de Ciencia de Datos utilizando metodologías ágiles.</li> <li>● Presentación de un plan de proyecto detallado, incluyendo asignación de recursos y cronograma.</li> <li>● Análisis de casos de estudio de proyectos para identificar lecciones aprendidas en la gestión de proyectos de Ciencia de Datos.</li> </ul>



## 8. Práctica(s)

- **Práctica de ética en ciencia de datos:**  
Los estudiantes trabajarán con conjuntos de datos reales y utilizarán herramientas como Python con bibliotecas como Pandas y Scikit-learn para analizar y mitigar sesgos algorítmicos. Utilizarán técnicas de preprocesamiento y postprocesamiento para garantizar la equidad y transparencia en los modelos desarrollados, además de discutir los resultados éticos de sus decisiones.
- **Práctica de Implementación de Modelos Avanzados:**  
Los alumnos implementarán una red neuronal convolucional utilizando TensorFlow en un entorno de computación en la nube como Google Colab. Utilizarán un conjunto de datos de imágenes para entrenar el modelo en la detección de objetos y luego explorarán técnicas de transferencia de aprendizaje para adaptar el modelo a un problema específico.
- **Práctica de Gestión de Proyectos de Ciencia de Datos:**  
En esta práctica, los estudiantes utilizarán herramientas de gestión de proyectos como Jira o Trello para planificar y coordinar un proyecto de ciencia de datos desde la definición del problema hasta la entrega del producto final. Trabajarán en equipos para asignar tareas, realizar seguimiento del progreso y gestionar eficazmente los recursos y el tiempo.

## 9. Proyecto de asignatura

El objetivo del proyecto que planteé el docente que imparta esta asignatura, es demostrar el desarrollo y alcance del(los) logro(s) formativo(s) de la asignatura, considerando las siguientes fases:

**Fundamentación:** marco referencial (teórico, conceptual, contextual, legal) en el cual se fundamenta el proyecto de acuerdo con un diagnóstico realizado, mismo que permite a los estudiantes lograr la comprensión de la realidad o situación objeto de estudio para definir un proceso de intervención o hacer el diseño de un modelo.

**Planeación:** con base en el diagnóstico en esta fase se realiza el diseño del proyecto por parte de los estudiantes con asesoría del docente; implica planificar un proceso de intervención empresarial, social o comunitario, el diseño de un modelo, entre otros, según el tipo de proyecto, las actividades a realizar los recursos requeridos y el cronograma de trabajo.

**Ejecución:** consiste en el desarrollo de la planeación del proyecto realizada por parte de los estudiantes con asesoría del docente, es decir en la intervención (social, empresarial), o construcción del modelo propuesto según el tipo de proyecto, es la fase de mayor duración que implica el desempeño de los saberes, habilidades y destrezas a desarrollar.

**Evaluación:** es la fase final que aplica un juicio de valor en el contexto laboral-profesión, social e investigativo, ésta se debe realizar a través del reconocimiento de logros y aspectos a mejorar se estará promoviendo el concepto de "evaluación para la mejora continua", el desarrollo del pensamiento crítico y reflexivo en los estudiantes.



## 10. Evaluación por competencias

La evaluación debe hacerse diagnóstica, formativa y sumativa. De igual manera, para fortalecer la parte actitudinal, se recomienda guiar al estudiante hacia la introspección para utilizar la autoevaluación y la coevaluación.

En el caso de las actividades de aprendizaje se sugiere el uso de estrategias metacognitivas como: mapas mentales, mapas conceptuales, reportes de prácticas, exposiciones en clase, ensayos, resúmenes, observación y cuestionarios, cuadros comparativos, informes.

Mientras que para verificar el nivel del logro de las competencias del estudiante se recomienda utilizar: el portafolio de evidencias, listas de cotejo, rúbricas, matrices de valoración, exámenes, guías de observación, además de estrategias en las que se logren las competencias blandas.

## 11. Fuentes de Información

1. Provost, F., & Fawcett, T. (2013). Data science for business: What you need to know about data mining and data-analytic thinking. O'Reilly Media, Inc.
2. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.
3. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media.
4. Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and machine learning. fairmlbook.org.
5. Murphy, K. P. (2012). Machine learning: a probabilistic perspective. MIT press.
6. Chollet, F. (2017). Deep learning with Python. Manning Publications.
7. O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Broadway Books.