



## 1. Datos Generales de la asignatura

<b>Nombre de la asignatura:</b>	Aprendizaje Automático
<b>Clave de la asignatura:</b>	TAD - 2501
<b>SATCA<sup>1</sup>:</b>	2-3-5
<b>Carrera:</b>	Ingeniería Informática

## 2. Presentación

### Caracterización de la asignatura

La asignatura Ciencia de Datos provee a los estudiantes de los conocimientos necesarios que le permitan identificar, formular y resolver problemas que requieran de modelos computacionales y soluciones tecnológicas para el procesamiento, almacenamiento y transferencia de información aplicando principios de ciencias de datos.

La inteligencia de negocios, el Machine Learning y el Análisis de Datos son elementos clave que impulsan, complementan y aceleran la estrategia de la empresa, apoyando el éxito de un nuevo modelo de negocio.

El análisis de datos y el aprendizaje automático permite mejorar las estrategias de la empresa. La asignatura ofrece herramientas para el análisis de datos afinando la capacidad de cualquier negocio para extraer y utilizar información de valor para construir ventajas competitivas.

Se incluyen temas de Deep Learning y Web Scraping que complementan una visión general de la Ciencia de Datos.

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





### Intención didáctica

El programa de la signatura consta de 4 temas, en el primer tema se presenta el Análisis de Datos donde el profesor proporciona una introducción a la materia así como las indicaciones para la instalación de las herramientas de soporte que se utilizarán durante el curso. Se exponen temas relacionados con el análisis de datos, librerías y el proceso del análisis de datos, a través de conceptos y ejercicios prácticos incluyendo la generación de archivos o lectura en formatos Excel y CSV y visualización de datos.

En el segundo tema, se expone el concepto de Machine Learning donde se desarrollan problemas de Regresión Lineal, Clasificación, Estadística Bayesiana, K Vecinos más próximos (KNN) y Árboles de Decisión.

En el tercer tema se expone una introducción a la Inteligencia Artificial y Redes Neuronales para posteriormente introducirse a Deep Learning. El Deep Learning lleva a cabo el proceso de Machine Learning usando una red neuronal artificial que se compone de un número de niveles jerárquicos para realizar aplicaciones de aprendizaje profundo utilizando herramientas de actualidad. El Deep Learning puede emplearse con éxito a grandes volúmenes de datos para el descubrimiento y aplicación de conocimiento, así como a la realización de predicciones a partir de él, principalmente debido a que obtiene tasas de éxito elevadas con entrenamiento “no supervisado”.

En el cuarto tema Web Scraping se incluyen las diferentes formas de obtener los datos, y solución a problemas básicos a través de distintas herramientas.

Los estudiantes deberán reforzar los temas vistos a través de la investigación y el autoaprendizaje y la realización de proyectos relacionados.

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

Lugar y fecha de elaboración o revisión	Participantes	Observaciones
---	---------------	---------------





Ciudad Guzmán, Jalisco, 3 de abril de 2025	Raquel Ochoa Ornelas	Programa elaborado para la conformación de la especialidad "Tecnologías Avanzadas para el Desarrollo de Software"
--	----------------------	---

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

##### Competencia(s) específica(s) de la asignatura

Desarrollar técnicas de ciencia de datos que permitan gestionar los datos en las empresas en modelos mediante el análisis predictivo implementando algoritmos de Machine Learning para mejorar la toma de decisiones de los negocios, así como implementar herramientas de Deep Learning y Web Scraping en el desarrollo de aplicaciones.

#### 5. Competencias previas

Conocimientos básicos de inteligencia de negocios para la toma de decisiones.

Disponer de conocimientos de programación y bases de datos para el desarrollo de aplicaciones.

Conocimientos básicos de matemáticas y estadísticas para la implementación de modelos.

#### 6. Temario

No.	Temas	Subtemas
1	Conceptos de Aprendizaje Automático y Modelos de Regresión	<b>1.1 Definición</b> <b>1.2 Tipos de aprendizaje</b> 1.2.1 Aprendizaje supervisado 1.2.2 Aprendizaje no supervisado 1.2.3 Aprendizaje por refuerzo 1.2.4 Aprendizaje semisupervisado y de transferencia <b>1.3 Modelos de regresión</b> 1.3.1 Regresión lineal 1.3.2 Regresión polinómica 1.3.3 Máquinas de soporte vectorial (SVR) para regresión





		1.3.4 Redes neuronales recurrentes (RNN) y LSTM para regresión
2	Modelos de Clasificación	Regresión logística Árboles de decisión Máquinas de soporte vectorial (SVM) K-Vecinos más cercanos (K-NN) Random Forest Naive-Bayes Redes neuronales artificiales (ANN) Redes neuronales convolucionales (CNN)
3	Evaluación de Modelos	<b>3.1 Métricas de Evaluación</b> 3.1.1 Regresión (MSE, $R^2$ , MAE, MAPE, RMSE) 3.1.2 Clasificación (Precisión, recall, F1-score, matriz de confusión, ROC-AUC). <b>3.2 Validación cruzada (Sobreajuste, rendimiento).</b> <b>3.3 Curvas de Aprendizaje</b> <b>3.4 Métodos de optimización</b> 3.4.1 Adam 3.4.2 RMSProp 3.4.3 Otros
4	Tendencias y Futuro de la Ciencia de Datos	<b>4.1 Inteligencia Artificial Explicable (XAI)</b> <b>4.2 Automatización de Modelos</b> <b>4.3 Inteligencia Artificial Generativa</b> <b>4.4 Inteligencia Artificial Responsable</b>





## 7. Actividades de aprendizaje de los temas

1. Análisis de datos	
Competencias	Actividades de aprendizaje
<p>Específica(s):</p> <ul style="list-style-type: none"><li>• Domina conceptos de ciencia de datos y los aplica a través de ejercicios prácticos.</li><li>• Realiza análisis exploratorio de datos de forma profesional.</li><li>• Genera estadísticas, efectúa operaciones de agrupamiento, visualiza gráficos y datos más complejos.</li><li>• Aprende a utilizar funciones estadísticas.</li></ul> <p>Genéricas:</p> <ul style="list-style-type: none"><li>• Sintetiza información encontrada en diferentes fuentes.</li><li>• Investiga de forma efectiva.</li><li>• Aprende de forma autónoma.</li></ul>	<ul style="list-style-type: none"><li>1.1 Instalar herramientas</li><li>1.2 Explorar el entorno, paquetes y canales</li><li>1.3 Desarrollar aplicaciones utilizando flujos de control</li><li>1.4 Implementar funciones en aplicaciones</li><li>1.5 Utilizar clases y estructuras de datos avanzadas</li><li>1.6 Aplicar distintos métodos de visualización de datos</li></ul>
2. Machine Learning	
Competencias	Actividades de aprendizaje





<p>Específica(s):</p> <ul style="list-style-type: none"> <li>• Procesa los datos para permitir su uso en Machine Learning.</li> <li>• Aplica diferentes tipos de algoritmos en Machine Learning.</li> </ul> <p>Genéricas:</p> <ul style="list-style-type: none"> <li>• Sintetiza información encontrada en diferentes fuentes.</li> <li>• Investiga de forma efectiva.</li> <li>• Aprende de forma autónoma.</li> </ul>	<ul style="list-style-type: none"> <li>• Definir un problema desde la perspectiva de Machine Learning</li> <li>• Despliega modelos en un sistema de producción</li> <li>• Decidir algoritmos a aplicar en un determinado problema</li> <li>• Generar estadísticas, efectúa operaciones de agrupamiento, visualiza gráficos y datos más complejos</li> <li>• Aprender a utilizar funciones estadísticas</li> </ul>
3. Deep Learning	
Competencias	Actividades de aprendizaje
<p>Específica(s):</p> <ul style="list-style-type: none"> <li>• Conoce y aplica técnicas para realizar una tarea a través del entrenamiento a una computadora.</li> </ul> <p>Genéricas:</p> <ul style="list-style-type: none"> <li>• Sintetiza información encontrada en diferentes fuentes.</li> <li>• Investiga de forma efectiva.</li> <li>• Aprende de forma autónoma.</li> </ul>	<p>3.1 Desarrollar una aplicación que permita evaluar el reconocimiento de una imagen</p>





4. Web Scraping	
Competencias	Actividades de aprendizaje
<p>Específica(s):</p> <ul style="list-style-type: none"> <li>Conoce y aplica técnicas para obtener información de páginas Web de forma automatizada.</li> </ul> <p>Genéricas:</p> <ul style="list-style-type: none"> <li>Sintetiza información encontrada en diferentes fuentes.</li> <li>Investiga de forma efectiva.</li> <li>Aprende de forma autónoma.</li> </ul>	<p>4.1 Agrupar datos de diferentes Webs para construir información conjunta</p> <p>4.2 Inspeccionar el código de una página Web para detectar cuáles son los elementos donde se incluyen los datos</p> <p>4.3 Generar alertas sobre cambios a una Web</p>

## 8. Práctica(s)

Práctica 1. Instalación del framework.  
 Práctica 2. Entornos, paquetes y canales.  
 Práctica 3. Introducción al lenguaje para el análisis de datos.  
 Práctica 4. Variables, estructuras de datos, flujos de control, funciones, clases, estructuras de datos avanzadas  
 Práctica 5. Argumentos en Script.  
 Práctica 6. Input-Output.  
 Práctica 7. Ejercicios con Numpy.  
 Práctica 8. Ejercicios con Pandas.  
 Práctica 9. Ejercicios de visualización.  
 Práctica 10. Ejercicios de análisis de datos para la ingesta de datos.  
 Práctica 11. Ejercicios de análisis de datos para el diagnóstico de la calidad de datos (QA).  
 Práctica 12. Ejercicios de análisis de datos para la agrupación de variables.  
 Práctica 13. Ejercicios de análisis de datos para la distribución de variables.  
 Práctica 14. Ejercicios de análisis de datos para comparaciones.







Práctica 15. Ejercicios de análisis de datos con herramientas adicionales.  
Práctica 16. Ejercicios de regresión lineal.  
Práctica 17. Ejercicio Naïve Bayes Classifier.  
Práctica 18. Ejercicio K vecinos más próximos.  
Práctica 19. Ejercicio con árboles de decisión.  
Práctica 20. Ejercicio descenso de Gradiente.  
Práctica 21. Ejercicio propagación hacia atrás.  
Práctica 22. Regularización en Deep Learning.  
Práctica 23. CNNs.  
Práctica 24. RNNs.  
Práctica 25. Scraping Sencillo.  
Práctica 26. Scraping Avanzado.  
Práctica 27. Scraping de JavaScript.







## 9. Proyecto de asignatura

Desarrollar un proyecto donde se apliquen los conocimientos adquiridos de la asignatura a través de un problema real detectado en el entorno, en una las siguientes vertientes:

- Salud
  - Procesos productivos
  - Procesos comerciales
  - Ciencias sociales
  - Comunicaciones
- **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.

La Ciencia de Datos (Data Science) se encarga de analizar grandes volúmenes de información con la ayuda de la inteligencia artificial. Además, se combinan técnicas de ciencia de la computación y la estadística, como el aprendizaje automático (machine learning), la inteligencia artificial, el reconocimiento de patrones, el aprendizaje estadístico, los modelos de probabilidad y la visualización, para administrar, analizar e interpretar datos.

Siendo los datos el valor más valorado de las empresas, con volúmenes históricos, el objetivo, es descubrir tendencias, pronósticos prediciendo el futuro, anticipándolo. Por lo anterior, las organizaciones requieren de profesionales que sepan manejar grandes cantidades de datos y herramientas de análisis. El mercado laboral informático necesita de profesionales con el conocimiento de ciencia de datos que puedan crear modelos capaces de predecir lo que ocurrirá en su sector.

El estudiante deberá visitar una empresa para detectar un problema relacionado con el análisis predictivo redactando una propuesta con los objetivos generales y objetivos específicos, métodos y herramientas del proyecto a utilizar. Se deben llevar a cabo





distintas encuestas para recabar información y estructurar los resultados que se esperan del proyecto.

- **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 las competencias genéricas y específicas a desarrollar. El estudiante describe la metodología y documenta el desarrollo de modelo a utilizar, realiza pruebas, analiza resultados parciales y detecta defectos para realizar correcciones.
- **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”, la metacognición, el desarrollo del pensamiento crítico y reflexivo en los estudiantes. El estudiante presenta los resultados de su modelo de probabilidad y la visualización de los datos.





## 10. Evaluación por competencias

Son las técnicas, instrumentos y herramientas sugeridas para constatar los desempeños académicos de las actividades de aprendizaje.

- Realizar ejercicios en las clases.
- Trabajos de investigación.
- Ejecutar cada una de las prácticas propuestas.
- Presentar examen teórico-práctico.
- Entrega de proyecto final.

## 11. Fuentes de información

1. García, J., Molina, J., Berlanga, A., Patricio, M., Bustamante, A., & Padilla, W. (2018). Ciencia de datos. *Técnicas Analíticas y Aprendizaje Estadístico*. Bogotá, Colombia. Publicaciones Altaria, SL.
2. Jones, H. (2019). *Ciencia de los datos: La guía definitiva sobre análisis de datos, minería de datos, almacenamiento de datos, visualización de datos, big data para empresas y aprendizaje automático para principiantes*.
3. Kelleher, J., Tierney, B. (2021). *Ciencia de datos*. Editorial: /Massachusetts Institute of Technology/Ediciones UC.
4. Lemus-Delgado, D., & Pérez Navarro, R. (2020). Ciencia de datos y estudios globales: aportaciones y desafíos metodológicos. *Colombia Internacional*, (102), 41-62.
5. Liberty, M. (2020). 50 principios de la ciencia de los datos. España, Barcelona. Editorial: BLUME.NET
6. Lima, P. B., Giana, G. E., Bollati, V. A., Pletsch, L. C., CInApTIC, U. T. N., & CONICET-CInApTIC, U. T. N. (2020). Ciencia de datos y Agilidad: una revisión de la literatura. *CibSE*, 222-235.
7. Rueda, R. A. S., Rueda, É. P. S., Rueda, R. D. S., & Pérez, Y. D. M. V. (2019). Análisis de la Aplicación Web Para la Estimación Puntual por medio de la Ciencia de Datos. *Dilemas contemporáneos: Educación, Política y Valores*.
8. Salas-Rueda, R. A., Salas-Rueda, É. P., & Salas-Rueda, R. D. (2019). Diseño y uso de una aplicación web para el campo de la estadística considerando el modelo Assure y la ciencia de datos. *Texto Livre: Linguagem e Tecnologia*, 12(1), 48-71.



9. Torres, J. (2020). Python Deep Learning: Introducción práctica con Keras y TensorFlow 2. Editorial: Marcombo.

