



1. Datos Generales de la asignatura

Nombre de la asignatura:	Análisis predictivo y machine learning
Clave de la asignatura:	ING-2509
SATCA¹:	3-3-6
Carreras:	Ingeniería en Sistemas Computacionales, Ingeniería Industrial, Ingeniería en Gestión Empresarial, Licenciatura en Administración, Ingeniería en Administración.

2. Presentación

Caracterización de la asignatura
<p>Formar profesionales de la ingeniería con las competencias necesarias para aplicar técnicas de análisis predictivo y aprendizaje automático en la resolución de problemas industriales, optimizando procesos, mejorando la toma de decisiones y promoviendo la innovación en el marco de la Industria 4.0.</p> <p>Se desarrollarán habilidades para recolectar, limpiar y preparar datos para sus análisis y aplicar modelos de aprendizaje automático adecuados a diferentes problemas. En diversas aplicaciones, se ha hecho necesario el desarrollar aplicaciones de análisis predictivo para la resolución de problemas reales en la industria.</p> <p>Además, permite gestionar proyectos de investigación y/o desarrollo tecnológico, así como ejercer actividades emprendedoras de liderazgo y adquirir habilidades para la toma de decisiones en su ámbito profesional.</p> <p>Para este propósito, se requiere de las competencias adquiridas en asignaturas que contengan temas como:</p> <ul style="list-style-type: none"> • Análisis matemático y optimización. • Instrumentación industrial. • Control y programación. <p>Su contenido impacta en las siguientes competencias del perfil de egreso.</p> <ul style="list-style-type: none"> • Desarrollo del pensamiento crítico y analítico para la resolución de problemas. • Trabajo en equipo y colaboración en proyectos de investigación. • Contribuye a la identificar las diferentes técnicas de aprendizaje supervisado, no supervisado y por refuerzo • Uso las herramientas y software más utilizados en el campo

¹ Sistema de Asignación y Transferencia de Créditos Académicos



Intención didáctica

La intención de esta asignatura es capacitar a los estudiantes para convertirse en analistas de datos capaces de extraer valor de grandes conjuntos de datos mediante el uso de técnicas de análisis predictivo y aprendizaje automático. A través de una combinación de teoría y práctica, los estudiantes adquirirán las habilidades necesarias para recolectar, limpiar, transformar y modelar datos, así como para interpretar los resultados y comunicar sus hallazgos de manera efectiva. Se busca fomentar un enfoque crítico y creativo en la resolución de problemas, incentivando a los estudiantes a explorar diferentes enfoques y a desarrollar soluciones innovadoras. La asignatura también hará énfasis en la importancia de la ética en el manejo de datos, sensibilizando a los estudiantes sobre las implicaciones sociales y éticas de sus decisiones. Al finalizar el curso, los estudiantes estarán preparados para enfrentar los desafíos del mundo laboral actual, donde la capacidad de analizar datos y extraer insights se ha convertido en una habilidad altamente demandada.

La metodología de la asignatura se centrará en la participación activa de los estudiantes. A través de proyectos prácticos, estudios de caso y trabajos colaborativos, los estudiantes tendrán la oportunidad de aplicar los conocimientos adquiridos en situaciones reales. Se utilizarán herramientas y software de última generación para facilitar el aprendizaje y el desarrollo de habilidades prácticas. Además, se organizarán seminarios y talleres con expertos en la materia para complementar la formación de los estudiantes y mantenerlos actualizados sobre las últimas tendencias en el campo. La evaluación de los estudiantes se basará en una combinación de actividades individuales y grupales, que incluirán exámenes teóricos, proyectos prácticos y presentaciones.

3. Participantes en el diseño y seguimiento curricular del programa

Lugar y fecha de elaboración o revisión	Participantes	Observaciones
Agosto-octubre 2024	Instituto Tecnológico de: Zitácuaro Huetamo Morelia Jiquilpan Lázaro Cárdenas La Piedad Uruapan	



4. Competencia(s) a desarrollar

Competencia(s) específica(s) de la asignatura
<ul style="list-style-type: none"> • Seleccionar y aplicar algoritmos de aprendizaje automático adecuados para resolver problemas específicos, considerando las características de los datos y los objetivos del análisis. • Realizar tareas de limpieza, transformación y selección de características en conjuntos de datos, utilizando técnicas como la eliminación de valores atípicos, la codificación de variables categóricas y la reducción de dimensionalidad. • Evaluar el desempeño de los modelos de aprendizaje automático utilizando métricas apropiadas, como precisión, recall, F1-score, y realizar ajustes a los hiperparámetros para mejorar su rendimiento. • Desarrollar aplicaciones de análisis predictivo para resolver problemas reales en diferentes dominios, como la industria, la salud, las finanzas, etc. Integrar los modelos de aprendizaje automático en sistemas más grandes y trabajar en colaboración con otros equipos para implementar soluciones.

5. Competencias previas

<ul style="list-style-type: none"> • Formula modelos matemáticos para la optimización de procesos. • Emplea la lógica algorítmica y lenguajes de programación. • Aplica los conceptos básicos de análisis de problemas complejos. • Manejo de bases de datos. • Conocimientos base en Big Data y la nube.
--

6. Temario

No.	Temas	Subtemas
1	Introducción al Análisis Predictivo y Machine Learning.	1.1. Conceptos básicos y aplicaciones del análisis predictivo. 1.2. Diferencias entre aprendizaje supervisado y no supervisado. 1.3. Ciclo de vida de un proyecto de machine learning.
2	Preparación de Datos para Modelado Predictivo	2.1. Selección y transformación de características. 2.2. División de datos en conjuntos de entrenamiento y prueba. 2.3. Técnicas de manejo de datos desbalanceados



3	Modelos de Aprendizaje Supervisado	<p>3.1. Protocolos Regresión lineal y logística.</p> <p>3.2. Almacén Árboles de decisión y bosque aleatorio.</p> <p>3.3. Seguridad Máquinas de soporte vectorial (SVM)</p>
4	Modelos de Aprendizaje No Supervisado	<p>4.1. Clustering y análisis de componentes principales (PCA)..</p> <p>4.2. Diseño Asociación y reducción de dimensionalidad.</p> <p>4.3. Aplicaciones prácticas de modelos no supervisados.</p>
5	Evaluación y Validación de Modelos	<p>5.1. Métricas de evaluación para modelos predictivos.</p> <p>5.2. Validación cruzada y ajuste de hiperparámetros.</p> <p>5.3. Técnicas de interpretación de modelos.</p>
6	Implementación y Despliegue de Modelos	<p>6.1. Integración de modelos en aplicaciones reales.</p> <p>6.2. Uso de APIs y servicios en la nube para despliegue.</p> <p>6.3. Mantenimiento y actualización de modelos en producción.</p>

7. Actividades de aprendizaje de los temas

1. Introducción al Análisis Predictivo y Machine Learning	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Tener un entendimiento de los conceptos clave, diferenciar entre los tipos de aprendizaje y asimilar las fases involucradas en un proyecto de machine learning.</p> <p><i>Genérica(s):</i></p> <ul style="list-style-type: none"> • Capacidad de análisis y síntesis. • Capacidad de investigación. • Solución de problemas. • Comunicación oral y escrita. 	<ul style="list-style-type: none"> • Crear una presentación o infografía que explique los conceptos básicos del análisis predictivo y el machine learning • Investigar y presentar un caso de estudio real de una empresa que haya utilizado el análisis predictivo para resolver un problema específico. • Desarrollar un ejercicio práctico en el que los estudiantes clasifiquen diferentes problemas de análisis de datos como de aprendizaje supervisado o no supervisado • Simular un proyecto de machine learning en grupos, asignando a cada estudiante un rol específico (recopilación de datos, limpieza, modelado, etc.).



2. Preparación de Datos para Modelado Predictivo	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Desarrollar competencias para aplicar técnicas de limpieza y preprocesamiento de datos, manejar datos faltantes y outliers, y realizar transformaciones y normalizaciones necesarias para asegurar la calidad de los datos.</p> <p><i>Genérica(s):</i></p> <ul style="list-style-type: none"> • Capacidad de análisis y síntesis. • Habilidad para buscar, procesar y analizar información procedente de fuentes diversas. • Capacidad de aplicar los conocimientos en la práctica. • Toma de decisiones. 	<ul style="list-style-type: none"> • Aplicar diferentes transformaciones a variables numéricas (normalización, estandarización) y categóricas (codificación one-hot, label encoding). • Explorar el impacto de diferentes transformaciones en el rendimiento de un modelo. • Taller Práctico de Limpieza de Datos: Organizar un taller donde los estudiantes trabajen con un conjunto de datos desordenado. Deberán aplicar técnicas de limpieza y preprocesamiento para corregir errores, estandarizar formatos y preparar los datos para su análisis. • Utilizar técnicas como la división aleatoria de datos. • Aplicar técnicas como el sobremuestreo (SMOTE), el submuestreo, el ajuste de pesos de clase y el uso de algoritmos específicos para datos desbalanceados
3. Modelos de Aprendizaje Supervisado	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Comprender los fundamentos de clasificación y regresión, así como los procesos de entrenamiento y evaluación.</p> <p><i>Genérica(s):</i></p> <ul style="list-style-type: none"> • Capacidad de análisis y síntesis. • Capacidad de aplicar los conocimientos en la práctica. • Habilidad de trabajar en forma autónoma. • Solución de problemas. • Habilidad para buscar y analizar información proveniente de fuentes diversas. 	<ul style="list-style-type: none"> • Implementar modelos de aprendizaje supervisado para resolver problemas reales, como la predicción de precios de viviendas, la clasificación de spam o la detección de fraudes. • Comparar el rendimiento de diferentes algoritmos en un mismo conjunto de datos y analizar los resultados. • Ajustar los hiperparámetros de los modelos para mejorar su rendimiento. • Estudiar casos reales de aplicación de modelos de aprendizaje supervisado en diferentes industrias.



4. Modelos de Aprendizaje No Supervisado	
Competencias	Actividades de aprendizaje
<p>Específica(s): Definir y diferenciar entre los principales problemas de aprendizaje no supervisado (clustering, reducción de dimensionalidad, detección de anomalías). Explicar cómo estos problemas se abordan sin la necesidad de datos etiquetados.</p> <p>Genérica(s):</p> <ul style="list-style-type: none"> • Capacidad de comunicación oral y escrita. • Capacidad de investigación. • Habilidades en el uso de las tecnologías de la información y comunicación Capacidad de investigación. • Toma de decisiones. 	<ul style="list-style-type: none"> • Aplicar algoritmos de aprendizaje no supervisado a conjuntos de datos reales, como la segmentación de clientes, la detección de fraudes o la compresión de imágenes. • Comparar el rendimiento de diferentes algoritmos de clustering en un mismo conjunto de datos y analizar los resultados. • Ajustar los hiperparámetros de los algoritmos para mejorar la calidad de los clusters o la reducción de la dimensionalidad. • Estudiar casos reales de aplicación de modelos de aprendizaje no supervisado en diferentes industrias.
5. Evaluación y Validación de Modelos	
Competencias	Actividades de aprendizaje
<p>Específica(s): Elegir las métricas de evaluación apropiadas según el tipo de problema (clasificación, regresión) y los objetivos del negocio. Calcular y analizar métricas como precisión, recall, F1-score, matriz de confusión, error cuadrático medio (MSE), raíz cuadrada del error cuadrático medio (RMSE), etc.</p> <p>Genérica(s):</p> <ul style="list-style-type: none"> • Capacidad de comunicación oral y escrita. • Capacidad de investigación. • Habilidades en el uso de las tecnologías de la información y comunicación Capacidad de investigación. • Toma de decisiones. 	<ul style="list-style-type: none"> • Proyectos prácticos: Evaluar y comparar el rendimiento de diferentes modelos en conjuntos de datos reales. • Experimentación con hiperparámetros: Ajustar los hiperparámetros de los modelos y analizar el impacto en el rendimiento. • Análisis de casos de estudio: Estudiar casos reales en los que la evaluación de modelos haya sido crucial para tomar decisiones importantes. • Desarrollo de pipelines de evaluación: Crear pipelines automatizados para evaluar modelos de manera eficiente. Crear planes de mantenimiento preventivo y correctivo para diferentes equipos.



6. Implementación y Despliegue de Modelos	
Competencias	Actividades de aprendizaje
<p>Específica(s): Crear interfaces de programación de aplicaciones (APIs) para exponer los modelos, permitiendo su integración con otras aplicaciones.</p> <p>Genérica(s):</p> <ul style="list-style-type: none"> • Capacidad de comunicación oral y escrita. • Capacidad de investigación. • Habilidades en el uso de las tecnologías de la información y comunicación Capacidad de investigación. <p>Toma de decisiones.</p>	<ul style="list-style-type: none"> • Proyectos prácticos: Desplegar modelos en diferentes entornos (local, nube) utilizando diferentes herramientas y tecnologías. • Desarrollo de pipelines de ML: Crear pipelines automatizados para el entrenamiento, evaluación y despliegue de modelos. • Experimentación con diferentes arquitecturas: Evaluar diferentes arquitecturas de microservicios y contenedores para desplegar modelos. • Análisis de casos de estudio: Estudiar casos reales de implementación de modelos en producción y los desafíos encontrados. • Participación en comunidades: Interactuar con otras personas en la comunidad de aprendizaje automático para compartir experiencias y aprender de los demás.

8. Práctica(s)

<ul style="list-style-type: none"> • Técnicas de limpieza y preprocesamiento: Aplicarán técnicas para limpiar y preprocesar los datos, asegurando su calidad. • Manejo de datos faltantes y outliers: Implementarán métodos para tratar datos faltantes y outliers, justificando sus decisiones. • Transformación y normalización: Realizarán transformaciones y normalizaciones necesarias para preparar los datos para el análisis. • Análisis de agrupamiento y segmentación: Implementarán algoritmos de agrupamiento para segmentar los datos y encontrar grupos significativos. • Reducción de dimensionalidad: Utilizarán técnicas como PCA para reducir la dimensionalidad de los datos y facilitar su análisis. • Uso de herramientas de software: Emplearán herramientas como Python o R para realizar exploraciones de datos complejas.
--



9. Proyecto de asignatura

El objetivo del proyecto que planteé el docente que imparta esta asignatura, es demostrar el desarrollo y alcance de la(s) competencia(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 las competencias genéricas y específicas 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”, la

10. Evaluación por competencias

Para evaluar las actividades de aprendizaje se recomienda solicitar: mapas conceptuales, reportes de investigación, estudios de casos, exposiciones en clase, reportes de visitas, portafolio de evidencias, tablas comparativas, glosarios, entre otros.

Para verificar el nivel del logro de las competencias del estudiante se recomienda utilizar: listas de cotejo, listas de verificación, matrices de valoración, rúbricas, entre otros, mediante la heteroevaluación, coevaluación, evaluación colectiva y autoevaluación.

En esta asignatura se requerirá la revisión del portafolio de evidencias, de las soluciones computarizadas, mapas, avances de proyecto, desarrollo de soluciones y el planteamiento de los análisis. En el contexto de:

Utilizar una rúbrica que contenga los niveles de alcance del estudiante:

- **Completo:** Que cubra todos los requisitos y que posea innovación y añadiduras de investigación.
- **Original:** Que sea una solución no copiada ni existente en internet u otras fuentes de conocimiento.
- **Funcional:** Que tenga las capacidades operativas mínimas.
- **Comprendido:** Que el estudiante conozca de fondo el proyecto y sea capaz de identificar todo lo realizado.



11. Fuentes de información

1. P. Norvig y S. J. Russell (2009), Artificial Intelligence: A Modern Approach, Third edition, Pearson.
2. Han, J., Kamber, M., & Pei, J. (2012). Data Mining: Concepts and Techniques.
3. Bagnato Juan I. (2020). Aprende Machine Learning (Teoría+Práctica Python).
4. Gerón, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly.
5. Harrison, M (2019). Machine Learning Pocket. Working with Structured Data in Python.
6. Mirjalili V., y Raschka S. (2019). Python Machine Learning. Marcombo.
7. Müller, A.C y Guido, S. (2016). Introduction to Machine Learning with Python: A Guide for Data Scientists. O'Reilly.