

1. Datos Generales de la asignatura

Nombre de la asignatura:	Análítica Inteligente de Datos.
Clave de la asignatura:	CDD-2105
SATCA¹:	2-3-5
Carrera:	Ingeniería En Sistemas Computacionales

2. Presentación

Caracterización de la asignatura

Las necesidades de la competitividad en el mercado hacen que las empresas privadas y las entidades de gestión pública estén interesadas en una explotación de sus bases de datos. La Minería de datos es el proceso de integración y tratamiento de los datos para convertirlos en información que permita apoyar a los tomadores de decisiones en la organización.

Esta asignatura aporta al perfil del Ingeniero en Sistemas Computacionales los conocimientos necesarios para el pre-proceso, análisis y modelización efectiva de los datos operativos e informacionales de la organización, mediante el uso de técnicas y herramientas estadísticas, así como de visualización.

Esta asignatura deberá cursarse en el módulo de especialidad, aproximadamente en el octavo semestre curricular de la carrera, teniendo como prerrequisito las materias de Almacenes de Datos y Grandes datos. Con la primera materia se relaciona con los temas de bases de datos para la toma de decisiones ya que el estudiante reforzará el conocimiento teórico y práctico relevante para el desarrollo de sistemas que faciliten la toma de decisiones. Con la materia Grandes datos se relaciona con la competencia específica que adquirirán en el uso de las herramientas para analizar y obtener conocimiento significativo que sea relevante para toma de decisiones en un ambiente corporativo.

Intención didáctica

El temario está organizado en cuatro temas; el primer tema se enfoca a conocer los conceptos y antecedentes de la Minería de Datos, los modelos que permiten obtener datos que indiquen si la empresa se encuentra en el margen de éxito o de pérdida, como son el modelo de verificación y el modelo de descubrimiento, los Problemas y resultados de la minería de datos. Así como, el proceso de minería de datos para realizar la preparación y procesamiento de los datos.

El tema dos, estudia el aprendizaje automático a través de las técnicas de clasificación

¹ Sistema de Asignación y Transferencia de Créditos Académicos

de la minería de datos para desarrollar aplicaciones que mejoren la competitividad de las empresas. Se trabaja con el aprendizaje supervisado tales como la clasificación con árboles de decisión, reglas de asociación, la clasificación con métodos Bayesianos, redes neuronales, y la regresión lineal simple, lineal múltiple y la regresión no lineal. Como parte de este tema se contempla el aprendizaje no supervisado debido a que los datos se pueden tratar como agrupaciones. Se contemplan los lenguajes R y Python para el desarrollo de las prácticas.

En el tema tres se estudia el contexto para la visualización de datos, las herramientas de análisis y visualización de datos que permiten conocer de manera óptima los resultados de los algoritmos del aprendizaje automático, otro de los temas que se abordará es el proceso de creación de una visualización de datos así como aspectos de la visualización de datos que puedan optimizar el proceso de toma de decisiones en los negocios de la organización como los reportadores, alertas, tableros de control e indicadores clave de desempeño.

3. Participantes en el diseño y seguimiento curricular del programa

Lugar y fecha de elaboración o revisión	Participantes	Observaciones
Instituto tecnológico de Zacatepec. Departamento de sistemas y Computación Academia de sistemas computación. Reunión para el desarrollo de Especialidades de Noviembre 2020 a Marzo 2021.	Tecnológico Nacional de México campus Instituto Tecnológico de Zactepec DR. Francisco Javier Cartujano Escobar M.C. Claudia Noguerón González Lic. Estela Rodríguez Zavaleta M.T.I. Laura Villavicencio Gómez M.C. Boris Antonio Aranda Benitez M.C. José Francisco Carpio Tovilla M.C.I. Misael López Nava M.C. Leticia Santa Olalla Ocampo Dra. Ana Celia Campos Hernández MT.I. Claudia Gabriela Bustillos Gaytan	Programa elaborado por profesores del departamento de Sistemas y Computación, presentado y aprobado en el pleno de la Academia. Esta materia forma parte de la especialidad Ciencia de Datos para la Toma de Decisiones , para la carrera Ingeniería en Sistemas Computacionales, Plan de estudios 2010

4. Competencia(s) a desarrollar

Competencia(s) específica(s) de la asignatura

Ser capaz de implementar las fases del proceso de descubrimiento de conocimiento a partir de bases de datos con las distintas técnicas de aprendizaje automático de acuerdo con el problema. Así como, generar los modelos y patrones elegidos utilizando una herramienta de minería de datos y visualización aplicado a casos de estudio reales.

5. Competencias previas

Dominar el diseño y desarrollo de Bases de Datos bajo un enfoque relacional.

Dominar el lenguaje estándar SQL para Bases de Datos.

Dominar programación básica en algún lenguaje de programación.

6. Temario

No.	Temas	Subtemas
1	Minería de datos	1.1 Conceptos de Minería de Datos. 1.2 Modelos de Minería de datos. 1.3 Problemas de la minería de datos. 1.4 Proceso de minería de datos
2	Aprendizaje automático	2.1 Clasificación con árboles de decisión 2.2 Clasificación con reglas de asociación 2.3 Clasificación con métodos Bayesianos 2.4 Clasificación con redes neuronales 2.5 Clasificación con regresión (lineal y no lineal) 2.6 Aprendizaje no supervisado
3	Visualización de datos	3.1 Contexto para la visualización de datos. 3.2 Herramientas de análisis y visualización de datos. 3.3 El proceso de creación de una visualización de datos. 3.4 Otros aspectos de la visualización de datos. 3.5 Reporteadores. 3.6 Alertas, tableros de control e indicadores clave de desempeño.

7. Actividades de aprendizaje de los temas

1. Minería de datos.	
Competencias	Actividades de aprendizaje
<p>Específica(s): Identificar los conceptos básicos, modelos y proceso de la minería de datos.</p> <p>Genéricas: Uso de las nuevas tecnologías, Comunicación oral y escrita, habilidades básicas de manejo de la computadora, trabajo en equipo, habilidades interpersonales, capacidad de análisis y síntesis, capacidad de organizar información, conocimientos básicos y avanzados del área de base de datos.</p>	<ul style="list-style-type: none"> • Profundizar en el concepto de Minería de Datos mediante ejercicios de cuestionamientos que puede resolver esta tecnología con algunas técnicas, así como el diseño y uso de herramientas que la aplican. • Realizar una investigación de la modelos de datos. • Identificar patrones válidos, novedosos, potencialmente útiles y, en última instancia, comprensibles a partir de los datos, teniendo como objetivo encontrar conocimiento útil relevante y nuevo sobre un fenómeno o actividad.
2. Aprendizaje automático	
Competencias	Actividades de aprendizaje
<p>Específica(s): Entender las técnicas que se deben aplicar para la generación de nuevo conocimiento.</p> <p>Genéricas: Uso de las nuevas tecnologías, Comunicación oral y escrita, habilidades básicas de manejo de la computadora, trabajo en equipo, habilidades interpersonales, capacidad de análisis y síntesis, capacidad de organizar información, conocimientos básicos y avanzados del área de base de datos.</p>	<ul style="list-style-type: none"> • Realizar un estudio del uso de herramientas que facilitan el análisis exploratorio de datos • Realizar ejercicios para cada técnica de minería de datos • Analizar datos a través de la clasificación • Aplicar el aprendizaje automático a un caso de estudio.
3. Visualización de datos	
Competencias	Actividades de aprendizaje

<p>Específica(s): Diseñar e implementar un sistema de visualización de datos.</p> <p>Genéricas: Uso de las nuevas tecnologías, Comunicación oral y escrita, habilidades básicas de manejo de la computadora, trabajo en equipo, habilidades interpersonales, capacidad de análisis y síntesis, capacidad de organizar información, conocimientos básicos y avanzados del área de base de datos.</p>	<ul style="list-style-type: none"> • Buscar y clasificar información sobre tecnologías y herramientas utilizadas para el análisis y visualización de datos • Elaborar prácticas para obtener resultados en las diferentes herramientas de visualización, tales como reportadores, indicadores clave de desempeño así como tableros de control.
---	--

8. Práctica(s)

<ul style="list-style-type: none"> • Realizar un análisis comparativo de diferentes herramientas para la minería de datos y para la Visualización de los datos. • Instalar y configurar la herramienta para el desarrollo de la materia. • Desarrollar una metodología que a partir de datos se obtenga información que ayude en la toma de decisiones. • Definir y aplicar esquemas de análisis de datos • Definir y aplicar algoritmos de minería de datos a utilizar • Realizar un estudio práctico de diferentes herramientas de apoyo a la toma de decisiones. • Realizar el análisis de datos con clasificación usando alguna herramienta de minería de datos.

9. Proyecto de asignatura

<p>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:</p> <ul style="list-style-type: none"> • 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 metacognición, el desarrollo del pensamiento crítico y reflexivo en los estudiantes.

10. Evaluación por competencias

La evaluación debe ser continua, formativa, flexible e integral, por lo que se debe considerar el desempeño en cada una de las actividades de aprendizaje, haciendo especial énfasis en lo siguiente:

- Realizar una evaluación diagnóstica al inicio del semestre, para partir de saberes previos, expectativas e intereses que tengan los estudiantes.
- Durante el desarrollo del curso debe llevarse a cabo una evaluación formativa que permita realimentar el proceso de aprendizaje y establecer las estrategias para el logro de los objetivos establecidos. Realizar evaluaciones mediante: exámenes teóricos y prácticos, desarrollo de prácticas de laboratorio, tarea y ejercicios.
- Al finalizar el curso debe realizarse una evaluación sumativa que se vincula con aquellas acciones que se orientan a dar cuenta de productos, saberes, desempeños y actitudes que se deben considerar para la calificación.
- Se sugiere utilizar como herramienta de evaluación el portafolio de evidencias, y como instrumento la lista de cotejo y la rúbrica.

Algunos productos sugeridos para la evaluación son:

- Reporte de investigaciones realizadas.
- Reporte de prácticas realizadas.
- Materiales utilizados en exposiciones.
- Proyecto integrador.
- Registros de la participación activa en la solución de estudio de casos.
- Examen escrito de conceptos teóricos y prácticos.

11. Fuentes de información

- Gareth James, D. W. (2017). *An Introduction to Statistical Learning: with Applications in R*. Springer.
- Hadley Wickham, G. G. (2017). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media Inc.
- Ian H. Witten, E. F. (n.d.). *Data Mining: Practical Machine Learning Tools and Techniques 4th Edition*. Morgan Kaufmann Series in Data Management Systems.
- McKinney, W. (2018). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython 2nd Edition*. O'Reilly Media Inc.
- Peter Bruce, A. B. (2020). *Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python 2nd Edition*. O'Reilly Media Inc.
- Vanderplas, J. (2017). *Python Data Science Handbook: Essential Tools for Working with Data*. O'Reilly Media Inc