



## Departamento de Matemáticas y Física

### Guía de Aprendizaje: Ciencia de Datos e Inteligencia de Negocios

#### DATOS GENERALES

<b>Asignatura:</b>	<b>Ciencia de Datos e Inteligencia de Negocios</b>			<b>Créditos BCD:</b>	<b>4</b>
<b>Clave:</b>	<b>ITE3073</b>	<b>Grupo:</b>		<b>Créditos TIE:</b>	<b>4</b>
<b>Horario:</b>	<b>9:00 – 11:00</b>	<b>Carreras:</b>	<b>Ingeniería Financiera</b>	<b>Salón:</b>	<b>En línea</b>
<b>Departamento:</b>	<b>Matemáticas y Física</b>	<b>UAB:</b>	<b>Matemáticas para Finanzas y Economía</b>	<b>Periodo:</b>	<b>Primavera 2021</b>
<b>Coordinador UAB:</b>	<b>Riemann Ruíz Cruz</b>		<b>E-mail:</b>	<b>RIEMANNRUIZ@iteso.mx</b>	<b>Ext. 3069</b>
<b>Profesor:</b>	<b>Gaddiel Desirena López</b>		<b>E-mail:</b>	<b>gdesirena@iteso.mx</b>	

BCD: *Bajo Conducción Docente*

TIE: *Trabajo Independiente del Estudiante*

#### PRESENTACIÓN

En el ámbito empresarial existe la búsqueda constante de eficiencia en los procesos empresariales para desarrollar una ventaja competitiva. De ahí surge la necesidad de implantar procesos de fabricación, de servicios o de logística que operen en condiciones óptimas. Las industrias continuamente se actualizan para mejorar sus procesos internos y así mejorar su eficiencia; pero para cada actualización se requiere hacer uso de información para evaluar o determinar las áreas de mejoras.

Las herramientas de cómputo, medición y/o bases de datos que dispone una industria, generan información que muchas veces no es analizada para la toma de decisiones. Actualmente, el volumen de datos que se genera por el uso de la tecnología es muy grande y la exploración de dichos datos para la obtención de información relevante para la toma de decisiones ha sido un tema que está teniendo mucho impacto.

La ciencia de datos es un campo interdisciplinario que involucra a los procesos y sistemas para extraer conocimiento o un mejor entendimiento de grandes volúmenes de datos en sus diferentes formas y formatos. Por otra parte, la inteligencia empresarial, inteligencia de negocios o BI (del inglés *business intelligence*), es el conjunto de estrategias, aplicaciones, datos, productos, tecnologías y arquitectura técnica, los cuales están enfocados a la administración y creación de conocimiento a través del análisis de los datos existentes en una organización o empresa.

La asignatura de Ciencia de Datos e Inteligencia de Negocios introduce técnicas de manejo y limpieza de datos para ser analizados y visualizados para la obtención de conocimiento y realizar la toma de decisiones en los negocios. Además, se plantea el fundamento, funcionamiento y la implementación de métodos de análisis de datos por medio de algoritmos de aprendizaje máquina (Machine Learning) supervisado y no supervisado.

## CONTEXTUALIZACIÓN DE LA ASIGNATURA

Para tener un desempeño satisfactorio de esta asignatura es indispensable contar con conocimientos de **programación, cálculo diferencial y estadística**. Es deseable manejar algún lenguaje científico de programación, por ejemplo, **MatLab, R, Python**.

Al finalizar con éxito la asignatura se tendrán las herramientas básicas para la limpieza, visualización de datos provenientes de diversas fuentes de información y el uso de métodos de clasificación para la obtención de conocimiento a partir de datos provenientes de fenómenos financieros como portafolios de inversión y series de tiempo de precios de activos financieros.

### MAPA DESCRIPTIVO



## PROPÓSITO GENERAL

Analizar bases de datos mediante el uso de algoritmos de aprendizaje máquina para la toma de decisiones en fenómenos financieros.

## PROPÓSITOS TRANSVERSALES

- a) Colaborar con los compañeros de clase.
- b) Generar el hábito de resolver problemas con herramientas computacionales.
- c) Usar de manera efectiva y eficaz las tecnologías de información para representar e interpretar los conceptos en diferentes formas: numérica, geométrica, algebraica y algorítmica.
- d) Implementar en Python los conocimientos adquiridos en este curso.

## ACTITUDES Y VALORES

Se espera que al cursar esta asignatura se desarrolle responsabilidad ante la actividad académica, manifiesta en al menos los siguientes aspectos:

- ☐ Participación activa, con compromiso, perseverancia y actitud positiva.
- ☐ El cumplimiento de las normas de disciplina establecidas.
- ☐ El cumplimiento en tiempo y forma de las actividades que se encomienden como trabajo independiente.
- ☐ El desarrollo de espíritu crítico y autocrítico (constructivo) en el análisis del desempeño propio y de los compañeros.
- ☐ El sentido de la ética, evitando, en particular, cometer actos deshonestos en la realización de las actividades evaluativas.
- ☐ El desarrollo de la capacidad para identificar características personales al afrontar procesos de aprendizaje y, como consecuencia, para aprender con mayor independencia.
- ☐ Diálogo abierto, directo y respetuoso tanto con el profesor como con los compañeros.
- ☐ Tolerancia y respeto

## PROPÓSITOS ESPECÍFICOS

Importar bases de datos provenientes de fuentes publicas

Seleccionar variables relevantes de una base de datos mediante el análisis estadístico de las mismas

Calcular parámetros estadísticos para describir la estructura de los datos

Visualizar información con la ayuda de gráficos y tablas

Clasificar datos por medio de algoritmos de clustering y clasificación supervisada.

## EVALUACIÓN DEL APRENDIZAJE

Productos	% de Calificación
Proyecto Final	40
3 exámenes parciales	30
Tareas	30
<b>TOTAL:</b>	<b>100%</b>

## **Políticas y lineamientos del curso.**

Tiempo Aproximado del Curso:

1. Descripción de datos (4 semanas)
2. Preprocesamiento de datos (5 semanas)
3. Aprendizaje Máquina (6 Semanas)
4. Proyecto (final) (1 semana)

- Las clases se impartirán en línea desde la plataforma Webex o Microsoft Teams.

- Cada módulo se evaluará con un examen teórico y un examen práctico de los conocimientos adquiridos en el módulo.
- Se realizarán prácticas durante las sesiones de clase para manejo de datos y obtención de resultados.
- Todas las prácticas serán realizadas en el software Python.
- Dentro de la sesión de clase no se permite hacer uso de celulares.
- El celular debe de estar en modo silencio o vibrador, de lo contrario el alumno tendrá que abandonar la sesión de clase.
- El pase de lista se hará en cada clase al **inicio de la sesión** y la cámara web prendida.
- Todas las tareas serán entregadas por medio de la plataforma **CANVAS** y **no se recibirán tareas a destiempo**.
- No se recibirán tareas tiempo después de la fecha límite; aunque se solicite que sea considerada con una calificación menor (Sean cuidadosos con la **organización de sus tiempos** y **cuidar los tiempos de entrega**).
- Durante la sesión de clase no se permite hacer desorden que afecte la impartición de la clase, los micrófonos deberán de estar apagados excepto si se les avisa lo contrario.
- Se requiere que el promedio de la calificación de los 3 exámenes parciales en el curso sea aprobatorio para tener derecho a calificación.
- Habiendo cumplido el criterio anterior, la calificación aprobatoria es **mayor o igual a 6.0** sin aplicar criterios de redondeo si se encuentra por debajo de este límite.
- Se permite como máximo el **20%** de inasistencias al curso, equivalente a **6 inasistencias** para mantener el derecho a calificación.

#### REFERENCIAS BIBLIOGRÁFICAS Y TELEMÁTICAS:

##### FUENTES REQUERIDAS

- Data Science for Business: What You Need to Know about Data Mining and Data-Analytic, Provost, F. y Fawcett, T., O'Reilly Media, 2013, ISBN: 978-1-449-37428-0 978-1-449-37429-7
- Foundations of Statistical Algorithms: with Reference to R Packages, Weihs, C. Mersmann, O. y Ligges, U., CRC Press, 2014, ISBN: 978-1-4398-7885-9.
- An Introduction to Machine Learning, Kubat, M., Springer, 2015, ISBN: 978-3-319-20009-5.
- Predictive Analytics with Microsoft Azure Machine Learning; Roger Barga, Valentine Fontama, Wee Hyong Tok, 2015, ISBN: 978-1-4842-1201-1
- Python Business Intelligence Cookbook; R. Dempsey, 2015, ISBN: 978-1-78528-966-8.

##### FUENTES DE APOYO

- El profesor proporcionará documentos de estudio con información relevante conforme el curso se vaya llevando a cabo. Este material se compartirá con los alumnos vía CANVAS.
- <https://librosweb.es/libro/python/>
- <http://pythonya.appspot.com/>
- <https://www.tensorflow.org>
- <http://www.mclibre.org/consultar/python/>

##### FUENTES DE DATOS

- <https://data.gov.uk/data/search>
- <https://datos.jalisco.gob.mx/>
- <https://www.kaggle.com/>
- <http://www.jaliscocomovamos.org/datos-abiertos>
- <https://datos.gob.mx/busca/dataset?organization=estado-de-jalisco>
- <https://datos.gob.mx/>
- <http://datos.guadalajara.gob.mx/>
- <http://www.transparencia.udg.mx/datos-abiertos>

- <http://datamx.io/dataset?q=bienevales>
- <https://finances.worldbank.org/all-datasets>
- <http://data.worldbank.org/>
- <https://www.data.gov/finance/>
- <http://www.kdnuggets.com/datasets/index.html>
- <https://www.treasury.gov/resource-center/financial-education/Pages/fdd.aspx>
- <http://www.datasciencecentral.com/profiles/blogs/big-data-sets-available-for-free>
- <https://web.stanford.edu/~hastie/ElemStatLearn/>