

1. DATOS DE LA ASIGNATURA

Nombre de la asignatura:	Analítica y Ciencia de Datos
Carrera:	Ingeniería en Tecnologías de la Información y Comunicaciones
Clave de la asignatura:	GTD-2101
SATC:	2-3-5

2. HISTORIA DEL PROGRAMA

Lugar y fecha de elaboración o revisión	Participantes	Observaciones (cambios y justificación)
Instituto Tecnológico de Culiacán, 30 de abril 2021.	Dr. Ricardo Rafael Quintero Meza M.C. Nora E. Cancela García	Creación de la Asignatura: Analítica y Ciencia de Datos 2021

3. PRESENTACIÓN

Caracterización de la asignatura

La toma de decisiones estratégica en las organizaciones depende principalmente de la capacidad que tiene para procesar toda la información que proviene de sus actividades operativas y de gestión, considerando los diversos eventos que suceden en las mismas con perspectivas del pasado, presente y futuro. La manera en que esta información se representa es muy diversa (estructurada, semiestructurada, no estructurada) por lo que su procesamiento demanda de múltiples herramientas (tecnológicas, matemáticas, estadísticas) para su análisis y explotación.

Más allá de las posibilidades que ofrecen las técnicas y herramientas tradicionales (almacenes de datos, inteligencia de negocios y otras que se gestaron en las décadas de los 80s y 90s del s. XX), la presencia de nuevos avances en la analítica de los datos y la ciencia de datos en este s. XXI abren posibilidades para la detección y aprovechamiento de hallazgos más amplios,

ofreciendo nuevas posibilidades para los tomadores de decisión de los niveles intermedios y estratégicos de las organizaciones.

La materia de Analítica y ciencia de datos pretende abordar estos temas emergentes como parte de la especialidad de TIC complementando de manera importante la formación de los futuros ITICs con un perfil que es altamente demandado por las organizaciones actuales. El curso cubre demandas de formación fundamental en programación, matemáticas, estadística e inteligencia artificial que se suelen requerirse para los científicos de datos.

Este curso se ofrece como materia optativa dentro de la especialidad

Intención Didáctica

La materia está diseñada bajo un enfoque basado en competencias que busca desarrollar principalmente, el saber, el saber hacer y el saber transferir con un nivel de desarrollo fundamental para el alumno, de tal manera que le permita abordar el tema en su campo laboral con bases para un desarrollo posterior en el mismo o bien tenga las bases para estudios posteriores a nivel de posgrado.

La intención didáctica de cada unidad es la siguiente:

En la Unidad 1 se introduce al estudiante en los conceptos, técnicas y herramientas para la analítica y la ciencia de datos, considerando como eje el lenguaje de programación Python, junto con una aproximación básica a las bibliotecas fundamentales para estas áreas de conocimiento: Numpy, Matplotlib y Pandas.

En la Unidad 2, se busca que el estudiante aborde los conceptos de la analítica de datos para representar lo que ha sucedido en la organización mediante la analítica descriptiva y la ciencia de datos. Para ello el estudiante hará uso de las herramientas Python y Pandas abordando las primeras fases del proceso de analítica de datos: extracción, limpieza y significado de los datos

En la Unidad 3, el estudiante abordará los conceptos de analítica de los datos y ciencia de datos para detectar hallazgos relevantes y de consideración para la organización a través de las herramientas Python y Pandas, así como de algunos modelos matemáticos fundamentales. Para ello practicará las fases del proceso de analítica de datos: graficación e interpretación de los datos.

En la Unidad 4, el estudiante se enfocará en la predicción abordando conceptos de analítica y ciencia de datos para encontrar patrones de comportamiento de los

datos y construir modelos matemáticos a partir de técnicas y herramientas del Machine Learning, haciendo uso del lenguaje Python y principalmente las bibliotecas Scikit-learn y StatsModels.

4. COMPETENCIAS A DESARROLLAR:

Competencia general:

Diseñar soluciones de analítica y ciencia de datos para fundamentar la toma de decisiones inteligente en las organizaciones

Competencias específicas:

Conocer los fundamentos de la programación para el diseño de herramientas de analítica y ciencia de datos.

Diseñar modelos analíticos y de ciencia de datos sobre lo que sucede en las organizaciones

Interpretar los modelos analíticos y de ciencia de datos para comprender las razones de lo que sucede en las organizaciones

Pronosticar lo que sucederá en las organizaciones a partir de modelos de analítica y ciencia de datos

Competencias genéricas:

Competencias instrumentales

- Capacidad de análisis y síntesis.
- Solución de problemas.
- Habilidad para buscar y analizar información proveniente de fuentes diversas.

Competencias interpersonales

- Capacidad crítica y autocrítica
- Trabajo en equipo
- Compromiso ético

Competencias sistémicas:

- Capacidad de aplicar los conocimientos en la práctica.
- Capacidad de aprender.

5. COMPETENCIAS PREVIAS

- Desarrolla soluciones de software para resolver problemas en diversos contextos utilizando acceso a datos, que soporten interfaz gráfica de

usuario y que ofrezcan soluciones a problemas del entorno.

- Resuelve problemas para apoyar la toma de decisiones aplicando técnicas matemáticas y estadísticas

6. TEMARIO

Unidad	Temas	Subtemas
1	Fundamentos de programación	1.1 Programación Python. 1.2 Paquetes Python para la analítica de datos (Numpy, Pandas, Matplotlib)
2	Analítica descriptiva de los datos	2.1 Introducción 2.2 Extracción y carga de datos 2.3 Limpieza de los datos 2.4 Significado de los datos
3	Analítica diagnóstica de los datos	3.1 Graficación de los datos 3.2 Interpretación de los datos
4	Analítica predictiva de los datos	4.1 Introducción 4.2 Machine Learning básico 4.3 Machine Learning avanzado

7. SUGERENCIAS DIDÁCTICAS PARA EL DESARROLLO DE COMPETENCIAS ESPECÍFICAS

Competencias específicas	Actividades de aprendizaje
1. Fundamentos de programación	Instalación de herramientas y ambiente de programación en Python.
Conocer los fundamentos de la programación para el diseño de herramientas de analítica y ciencia de datos.	Revisión y práctica de conceptos de programación básica en Python. Revisión y práctica de conceptos de bibliotecas Numpy, Pandas y Matplotlib de Python.

<p>2. Analítica descriptiva de los datos</p> <p>Diseñar modelos analíticos y de ciencia de datos sobre lo que sucede en las organizaciones</p>	<p>Revisión y práctica de conceptos de extracción y limpieza de datos con la biblioteca Pandas y Python.</p> <p>Revisión y práctica de conceptos para la determinación de significado de los datos haciendo uso de modelos matemáticos y estadísticos y las facilidades que ofrece la biblioteca Pandas y Python</p>
<p>3. Analítica diagnóstica de los datos</p> <p>Interpretar los modelos analíticos y de ciencia de datos para comprender las razones de lo que sucede en las organizaciones</p>	<p>Revisión y práctica de conceptos matemáticos y estadísticos para la detección e interpretación de los datos haciendo uso de las facilidades de la biblioteca Pandas y Python.</p>
<p>4. Analítica predictiva de los datos</p> <p>Pronosticar lo que sucederá en las organizaciones a partir de modelos de analítica y ciencia de datos</p>	<p>Revisión y práctica de conceptos matemáticos y de Inteligencia Artificial para la construcción de modelos de predictivos de los datos haciendo uso de las bibliotecas Scikit-Learn, StatsModels y Python</p>

8. SUGERENCIAS DIDÁCTICAS PARA EL DESARROLLO DE COMPETENCIAS GENÉRICAS.

<ol style="list-style-type: none"> 1. Ejercicios de programación básicos sobre las construcciones fundamentales del lenguaje Python: entrada-salida, estatutos condicionales, estatutos cíclicos, datos primitivos, colecciones, funciones y módulos. 2. Ejercicios de programación para practicas las bibliotecas base para la analítica y ciencia de datos: arreglos, matrices, dataframes y graficación. 3. Ejercicios de programación basados en la creación de Jupyter Notebooks para ejercitar los diversos casos de uso de la analítica descriptiva basados en la

biblioteca Pandas. Incluir la publicación de resultados mediante tecnologías en la nube disponibles en servicios como Github, Gist y otros.

4. Ejercicios de programación basados en la creación de Jupyter Notebooks para ejercitar los diversos casos de uso de la analítica diagnóstica basados en la biblioteca Pandas. Incluir la publicación de resultados mediante tecnologías en la nube disponibles en servicios como Github, Gist entre otros.
5. Ejercicios de programación basados en la creación de Jupyter Notebooks para ejercitar los diversos casos de uso de la analítica prescriptiva basados en las bibliotecas Scikit-Learn y StatsModels. Incluir la publicación de resultados mediante tecnologías en la nube disponibles en servicios como Github, Githist y otros.

9. SUGERENCIAS DE EVALUACIÓN

Diagnóstica	Formativa	Sumativa
Aplicación de cuestionario sobre conocimientos previos de programación y análisis numérico.	Aplicación de cuestionario sobre temas estudiados. Presentación de avances del proyecto de la materia	Examen teórico-práctico de conocimientos. Evaluación de avances del proyecto de la materia.

10. FUENTES DE INFORMACIÓN

Libros de texto

- Henley A.J., Wolf D. Learn Data Analysis with Python. Apress. 2018.
- Swamynathan M. Mastering Machine Learning with Python in Six Steps. Apress. 2019.

11. PERFIL DEL PROFESOR QUE IMPARTIRÁ LA MATERIA

- Doctor en Ciencias Computacionales con experiencia práctica en desarrollo de aplicaciones Ciencia de Datos.
- Maestro en Inteligencia Artificial ó Maestro en Ciencias Computacionales con experiencia práctica en desarrollo de desarrollo de aplicaciones Ciencia de Datos.
- Ingeniero en Sistemas Computacionales, Licenciado en Informática o carrera afín con experiencia desarrollo de aplicaciones Ciencia de Datos.