

1. Datos Generales de la asignatura

Nombre de la asignatura:	<i>Fundamentos de Análisis de Datos con Python.</i>
Clave de la asignatura:	<i>DWD-2105</i>
SATCA¹:	<i>(2-3-5)</i>
Carrera:	<i>Ingeniería en Sistemas Computacionales.</i>

2. Presentación

Caracterización de la asignatura

*Python es un lenguaje de programación de tipo general que soporta un desarrollo rápido de aplicaciones de análisis de datos. La palabra “Python” es usada para referirse tanto al lenguaje de programación, como a la herramienta que ejecuta los **scripts** escritos en el lenguaje “Python”.*

Sus principales ventajas son:

- Es gratis.*
- De código abierto.*
- Disponible para todas las plataformas más importantes (macOS, Linux, Windows).*
- Es mantenido por la Python Software Foundation.*
- Soporta múltiples paradigmas de programación.*
- Tiene una gran comunidad.*
- Tiene un rico ecosistema de paquetes de terceros.*

Por otro lado, el análisis de datos es la ciencia que se encarga de examinar un conjunto de datos con el propósito de sacar conclusiones sobre la información para tomar decisiones, o simplemente ampliar los conocimientos sobre diversos temas.

Python provee un marco de trabajo que permite combinar enfoques de diferentes disciplinas de investigación (y no solo de investigación) para análisis.

Esta asignatura proporciona al estudiante los conocimientos básicos para conocer y hacer uso de Python para resolver problemas de aplicaciones de la vida cotidiana y de la ingeniería utilizando la analítica de datos. Está diseñada para el logro de cuatro competencias específicas dirigidas a la comprensión de los dominios de: Introducción, Python y Pandas, Análisis de Datos y Programación con Python para IoT.

¹ Sistema de Asignación y Transferencia de Créditos Académicos

Intención didáctica

La asignatura pretende proporcionar al alumno conocimientos en el análisis de datos con Python y algunas bases para aplicar Python en otras áreas.

Se organiza el contenido de la materia en cuatro temas.

El primer tema provee una visión general del curso y los tópicos que serán abordados.

En el segundo tema se introduce al estudiante a Python y a las estructuras que maneja.

En el tercer tema se introduce al estudiante al área de machine learning basada en Python y sus librerías.

Y, por último, en el cuarto tema se abordan temas diferentes al análisis de datos donde se pueda usar Python como lenguaje de propósito general.

Se proponen actividades de aprendizaje que permitan al estudiante el desarrollo de las competencias requeridas, y se propone adecuarlas al contexto.

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.	M.C. Luis José Muñiz Rascado. Dr. Jesús Ángel Peña Ramírez, M.C. Norma Josefina Ontiveros Hernández, Dr. Sócrates Espinoza Salgado, M.C. Mario Humberto Tiburcio Zúñiga, Ing. Alejandra Calypso Santa Olalla Salgado.	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 Desarrollo Web Empresarial , para la carrera Ingeniería en Sistemas Computacionales, Plan de estudios 2010

4. Competencia(s) a desarrollar

Competencia(s) específica(s) de la asignatura

Aplica la analítica de datos para obtener conocimiento, usando como herramienta el lenguaje *Python* y sus librerías, a partir de datos de diferentes áreas para la toma de decisiones.

5. Competencias previas

- Elabora programas en al menos un lenguaje de programación para asimilar los nuevos paradigmas en *Python*.
- Diseña e implementa bases de datos en al menos un manejador de base de datos para incursionar en el análisis de repositorios.
- Conoce y utiliza los conceptos de estadística para tomar decisiones con determinada verosimilitud para el análisis de repositorios datos.
- Comprende y aplica la teoría de conjuntos para determinar la pertenencia de objetos en determinadas clases.

6. Temario

No.	Temas	Subtemas
1	Introducción.	1.1 Introducción a <i>Python</i> . 1.2 Introducción al análisis de datos. 1.3 Ciencias de datos. 1.4 <i>Machine learning</i> . 1.5 <i>Deep learning</i> . 1.6 Lenguajes y herramientas de análisis de datos.
2	Python y Pandas.	2.1 Listas, tuplas y diccionarios. 2.2 Condicionales y ciclos. 2.3 Manejo de archivos. 2.4 Introducción a Pandas. 2.5 Pandas Series. 2.6 Pandas <i>Dataframes</i> . 2.7 Conexión a base de base de datos (<i>NoSQL</i> , <i>RDBMS</i>) con Pandas. 2.8 Filtrado de datos. 2.9 Transformación de datos. 2.10 <i>Discretización y bucketización</i> . 2.11 Expresiones regulares. 2.12 <i>Join y Merge</i> . 2.13 Operaciones de pivote. 2.14 Funciones de agregación. 2.15 Visualización y ploteo.
3	Análisis de Datos.	3.1 Estadística básica. 3.1.1 Funciones estadísticas. 3.2 Series de tiempo. 3.3 Análisis supervisado.

		3.3.1 Regresión Lineal Simple. 3.3.2 Regresión Lineal Múltiple. 3.3.3 KNN. 3.3.4 Árboles de Decisión. 3.3.5 <i>Random Forest</i> . 3.3.6 Redes Neuronales y Deep Learning. 3.3.7 Máquinas de Soporte Vectorial (SVM). 3.4 Análisis no supervisado. 3.4.1 Métodos de Agrupamiento. 3.4.1.1 <i>K-Means</i> . 3.4.1.2 DBSCAN.
4	Programación con <i>Python</i> para <i>IoT</i> .	4.1 Concepto de <i>IoT</i> . 4.2 Programación Python básica para <i>IoT</i> . 4.3 <i>Framework IoT</i> . 4.4 Manejo de sensores, actuadores y módulos electrónicos con la placa electrónica básica de <i>IoT</i> . 4.5 Comunicaciones y protocolos. 4.6 Plataformas en la nube.

7. Actividades de aprendizaje de los temas

Nombre de tema Introducción	
Competencias	Actividades de aprendizaje
Específica(s): Obtiene una visión general del contenido y objetivos del curso sobre el uso de <i>Python</i> como herramienta de análisis de datos. Genéricas: <ul style="list-style-type: none"> - Habilidad de investigación. - Capacidad de aplicar los conocimientos en la práctica. - Capacidad de generar nuevas ideas. - Optimiza soluciones. - Usa el pensamiento lógico, algorítmico, heurístico, analítico y sintético. 	<ul style="list-style-type: none"> * Realizar una búsqueda y revisión de los conceptos de <i>machine learning</i>, y de los alcances de las herramientas de análisis de datos, por equipos y consultando diversas fuentes de información. * Realizar prácticas de laboratorio.
Nombre de tema <i>Python</i> y <i>Pandas</i> .	
Competencias	Actividades de aprendizaje

<p>Específica(s):</p> <p>Familiariza al estudiante con el lenguaje Python y las principales estructuras de datos que éste provee como preámbulo para el aprendizaje posterior de técnicas de <i>machine learning</i>.</p> <p>Genéricas:</p> <ul style="list-style-type: none"> - Habilidad para buscar y analizar. - Trabajo en equipo. - Habilidad de investigación. - Capacidad para aprender. - Capacidad de análisis y síntesis. - Capacidad crítica y autocrítica. - Habilidad de investigación. 	<ul style="list-style-type: none"> * Realizar diferentes ejercicios sobre el uso y manejo de estructuras de datos Numpy y Pandas, tales como creación de estructuras, operaciones aritméticas entre vectores y vectores y escalares que ilustren la vectorización, Así como despliegue de datos, agregar datos a una estructura.
<p>Nombre de tema Análisis de Datos.</p>	
<p>Competencias</p>	<p>Actividades de aprendizaje</p>
<p>Específica(s):</p> <p>Analiza datos utilizando técnicas de <i>machine learning</i> en repositorios de datos.</p> <p>Genéricas:</p> <ul style="list-style-type: none"> - Capacidad de análisis y síntesis. - Habilidad para buscar y analizar. - Trabajo en equipo. - Capacidad crítica y autocrítica. - Habilidad de investigación. - Capacidad para aprender. - Capacidad de aplicar los conocimientos en la práctica. 	<ul style="list-style-type: none"> * Realizar diferentes ejercicios de modelado y análisis de datos donde puedan conocer las librerías que integran las técnicas de <i>machine learning</i>. * Realizar la secuencia de análisis de datos de análisis de datos que va desde lectura y limpieza de datos, el entrenamiento de los modelos, así como su aplicación a los datos de prueba y el análisis de los resultados que incluya la revisión de la matriz de confusión y precisión aportada por cada técnica.
<p>Nombre de tema Programación con Python para <i>IoT</i>.</p>	
<p>Competencias</p>	<p>Actividades de aprendizaje</p>

<p>Específica(s):</p> <p>Envía datos de sensores a la nube y realiza servicios remotos siguiendo los conceptos de <i>IoT</i> para la operación de un tablero de control.</p> <p>Genéricas:</p> <ul style="list-style-type: none"> - Capacidad de análisis y síntesis. - Habilidad para buscar y analizar. - Trabajo en equipo. - Capacidad de aplicar los conocimientos en la práctica. 	<ul style="list-style-type: none"> * Realizar prácticas para enviar datos de sensores a la nube. * Diseñar un tablero de control (<i>dashboard</i>). * Realizar un servicio <i>IoT</i>.
---	--

8. Práctica(s)

<ul style="list-style-type: none"> • Entrar a <i>Google Colab</i> y determinar su funcionamiento a través de la realización de ejercicios. • Descargar e instalar Anaconda para trabajar <i>Python</i> con <i>Spyder</i>. • Análisis de datos usando técnicas de <i>machine learning</i> con <i>Python</i> siguiendo la siguiente secuencia para cada técnica: <ol style="list-style-type: none"> 1. Importar conjunto de datos. 2. Limpiar y preparar datos para análisis. 3. Dividir datos en dos conjuntos: entrenamiento y prueba. 4. Entrenar un modelo de <i>machine learning</i> usando <i>scikit-learn</i>, para cada metodología de <i>machine learning</i> del temario. 5. Aplicar el modelo entrenado a los datos de prueba. 6. Calcular y analizar la precisión del modelo. 7. Determinar, imprimir y graficar la importancia de las características utilizadas en el análisis. • Realizar los siguientes tres ejercicios: <ol style="list-style-type: none"> a. Comunicaciones <i>IoT</i> a la nube (MQTT) b. Diseño de un <i>dashboard</i>. c. Realización de un servicio <i>IoT</i>
--

9. Proyecto de asignatura

--

Buscar en Internet un conjunto de datos y analizarlos usando técnicas de *machine learning* de tal modo que puedan derivar o predecir algún conocimiento de esos datos.

- **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

Tema 1:

- Cuadro sinóptico
- Rúbrica del trabajo Grupal.
- Instalar y usar

Tema 2:

- Metodología de trabajo.
- Rúbrica del trabajo grupal.
- Desenvolvimiento en el desarrollo de la práctica de laboratorio.
- Informe de la práctica de laboratorio.

Tema 3:

- Cuestionario.
- Rúbrica del trabajo grupal
- Desenvolvimiento en el desarrollo de la práctica de laboratorio.



- Informe de la práctica de laboratorio.

Tema 4:

- Elaboración de Tabla.
- Rúbrica del trabajo grupal
- Desenvolvimiento en el desarrollo de la práctica de laboratorio.
- Informe de la práctica de laboratorio.

11. Fuentes de información

1. Comunidad Python.
2. Wilkinson C. (2020). Ciencia de Datos Python. Estados Unidos: Independently Published.
3. Hush J. (2020). Formato Libro físico Autor John Hush Editorial Lightning Source Inc Isbn13 9781801136143 Compartir Python Para el Análisis de Datos. Estados Unidos: Lightning Source Inc.
4. Date C. J., (2009). Introducción a los Sistemas de Bases de Datos, 7a. edición: Prentice Hall.
5. Silberschatz, Korth, Sudarshan, (2006). Fundamentos de Bases de Datos, 4ª. Edición, Mc Graw Hill. Direcciones electrónicas:
<https://iaarbook.github.io/datascience/>