



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



PROGRAMA SINTÉTICO

UNIDAD ACADÉMICA: UNIDAD PROFESIONAL INTERDISCIPLINARIA EN INGENIERÍA CAMPUS COAHUILA
ESCUELA SUPERIOR DE CÓMPUTO
UNIDAD PROFESIONAL INTERDISCIPLINARIA EN INGENIERÍA CAMPUS TLAXCALA

PROGRAMA ACADÉMICO: Licenciatura en Ciencia de Datos

UNIDAD DE APRENDIZAJE: Desarrollo de Aplicaciones para análisis de Datos

SEMESTRE: IV

PROPÓSITO DE LA UNIDAD DE APRENDIZAJE:

Desarrolla aplicaciones de análisis de datos a partir de su adquisición, preprocesamiento, y extracción de conocimiento.

| | | | | |
|-----------------------------------|---|---|---|---|
| CONTENIDOS: | I. Programación con Lenguaje Python II. Adquisición y almacenamiento de Datos III. Aprendizaje automático | | | |
| ORIENTACIÓN DIDÁCTICA: | Métodos de enseñanza | | Estrategias de aprendizaje | |
| | a) Deductivo | X | a) Estudio de Casos | |
| | b) Inductivo | X | b) Aprendizaje Basado en Problemas | X |
| | c) Analógico | X | c) Aprendizaje Orientado a Proyectos | |
| | e) Especializado | | e) | |
| | f) Heurístico | X | | |
| EVALUACIÓN Y ACREDITACIÓN: | Diagnóstica | X | Saberes Previamente Adquiridos | X |
| | Solución de casos | X | Organizadores gráficos | X |
| | Problemas resueltos | X | Problemarios | |
| | Reporte de proyectos | | Reporte de seminarios | |
| | Reportes de indagación | | Otras evidencias a evaluar: Discusión dirigida | |
| | Reportes de prácticas | X | Programas de cómputo utilizando el lenguaje Python y sus bibliotecas. | |
| | Evaluaciones escritas | | | |

| BIBLIOGRAFÍA BÁSICA: | Autor(es) | Año | Título del documento | Editorial / ISBN |
|-----------------------------|------------------|------------|--|--------------------------|
| | Embarak, O. | 2018 | <i>Data Analysis and Visualization Using Python</i> | Apress / 9781484241080 |
| | László, G. | 2018 | <i>Website Scraping with Python</i> | Apress / 9781484239247 |
| | McKinney, M. | 2018 | <i>Python for Data Analysis</i> | O'Reilly / 9781491957660 |
| | Nelly, F. | 2018 | <i>Python Data Analytics</i> | Apress / 9781484239124 |
| | Porcu, V. | 2018 | <i>Python for Data Mining Quick Syntax Reference</i> | Apress / 9781484241127 |



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



PROGRAMA DE ESTUDIOS

UNIDAD DE APRENDIZAJE: Desarrollo de Aplicaciones para análisis de Datos

HOJA 2 DE 7

UNIDAD ACADÉMICA: ESCUELA SUPERIOR DE CÓMPUTO, UNIDAD PROFESIONAL INTERDISCIPLINARIA EN INGENIERÍA, CAMPUS COAHUILA

PROGRAMA ACADÉMICO: Licenciatura en Ciencia de Datos

| | | |
|---|--|-----------------------------------|
| SEMESTRE: IV | ÁREA DE FORMACIÓN: Profesional | MODALIDAD: Escolarizada |
| TIPO DE UNIDAD DE APRENDIZAJE: Teórica-Práctica/Obligatoria | | |
| VIGENTE A PARTIR DE: Agosto, 2021 | CRÉDITOS: Tepic: 7.5 | SATCA: 6.1 |

INTENCIÓN EDUCATIVA

Esta unidad contribuye al perfil de egreso de la Licenciatura en Ciencia de Datos con el desarrollo de habilidades básicas de programación, abstracción, pensamiento analítico y sistemático en Lenguaje Python para la elaboración de bibliotecas especializadas en la adquisición de datos desde la web, su preprocesamiento, almacenamiento de datos y extracción de conocimiento con métodos de aprendizaje automático. Así mismo fomenta el trabajo en equipo, resolución de problemas, comunicación efectiva, creatividad e ingenio.

La presente unidad se relaciona de manera antecedente con Programación para Ciencia de Datos y Bases de Datos, de manera lateral con Bases de Datos Avanzadas y Desarrollo de Aplicaciones Web y de manera consecuente con Analítica y Visualización de Datos, Aprendizaje de Máquina e Inteligencia Artificial y Minería de Datos.

PROPOSITO DE LA UNIDAD DE APRENDIZAJE

Desarrolla aplicaciones de análisis de datos a partir de su adquisición, preprocesamiento, y extracción de conocimiento.

| | | |
|--|---|--|
| TIEMPOS ASIGNADOS | UNIDAD DE APRENDIZAJE DISEÑADA POR: Comisión de Diseño del Programa Académico. | AUTORIZADO Y VALIDADO POR: |
| HORAS TEORÍA/SEMANA: 3.0 HORAS PRÁCTICA/SEMANA: 1.5 HORAS TEORÍA/SEMESTRE: 54.0 HORAS PRÁCTICA/SEMESTRE: 27.0 HORAS APRENDIZAJE AUTÓNOMO: 20.0 HORAS TOTALES/SEMESTRE: 81.0 | APROBADO POR: Comisión de Programas Académicos del H. Consejo General Consultivo del IPN. 22/10/2020 | Ing. Juan Manuel Velázquez Peto Director de Educación Superior |



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Desarrollo de Aplicaciones para análisis de Datos

HOJA 3 **DE** 7

| UNIDAD TEMÁTICA I Programación con lenguaje Python | CONTENIDO | HORAS CON DOCENTE | | HR S AA |
|---|---|--------------------------|----------|----------------|
| | | T | P | |
| UNIDAD DE COMPETENCIA Analiza la programación de computadoras a partir de sus fundamentos y el lenguaje Python. | 1.1 Aspectos básicos | 1.0 | 1.0 | 1.0 |
| | 1.2 Tipos de datos estándar en Python | 1.0 | 1.0 | 1.0 |
| | 1.3 Estructuras de control | 3.0 | 1.0 | 1.0 |
| | 1.4 Funciones | 3.0 | 1.0 | 1.0 |
| | 1.5 Programación Orientada a Objetos con Python | 10.0 | 5.0 | 3.0 |
| | Subtotal | 18.0 | 9.0 | 7.0 |

| UNIDAD TEMÁTICA II Adquisición y almacenamiento de Datos | CONTENIDO | HORAS CON DOCENTE | | HR S AA |
|---|---|--------------------------|----------|----------------|
| | | T | P | |
| UNIDAD DE COMPETENCIA Desarrolla programas en lenguaje Python en la adquisición de datos de diferentes orígenes y tipos, y su preprocesamiento. | 2.1 Datos estructurados y semiestructurados | 2.0 | 1.0 | 1.0 |
| | 2.1.1 Archivos .csv y .tsv | | | |
| | 2.1.2 Archivos binarios HDF5 y Excel | | | |
| | 2.1.2 JSON | | | |
| | 2.1.3 XML | | | |
| | 2.2 Almacenamiento de datos | 2.0 | 2.0 | 1.0 |
| | 2.2.1 Base de datos relacional | | | |
| | 2.2.2 Base de datos NoSQL | | | |
| | 2.3 Bibliotecas de Python para estructurar datos | 4.0 | 2.0 | 1.0 |
| | 2.3.1 NumPy | | | |
| | 2.3.2 Pandas | | | |
| | 2.3.3 SciPy | 6.0 | 2.0 | 2.0 |
| | 2.4 Web Scraping | | | |
| | 2.4.1 Biblioteca Request | | | |
| | 2.4.2 Biblioteca BeautifulSoup | | | |
| | 2.4.3 Biblioteca Scrapy | | | |
| | 2.4.4 Biblioteca Selenium | | | |
| | 2.4.5 Biblioteca PythonAnywhere | | | |
| | 2.4.6 Recolección de datos de Redes Sociales | | | |
| | 2.4.7 Recolección de datos de geolocalización | | | |
| | 2.5 Preprocesamiento de datos | 4.0 | 2.0 | 1.0 |
| | 2.5.1 Limpieza de datos y manejo de datos faltantes | | | |
| | 2.5.2 Transformación de Datos | | | |
| | 2.5.3 Manipulación de cadenas | | | |
| | 2.5.4 Volcado de datos | | | |
| | Subtotal | 18.0 | 9.0 | 6.0 |



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Desarrollo de Aplicaciones para análisis de Datos

HOJA 4 **DE** 7

| UNIDAD TEMÁTICA III Aprendizaje automático | CONTENIDO | HORAS CON DOCENTE | | HR S AA |
|---|---|--------------------------|----------|----------------|
| | | T | P | |
| UNIDAD DE COMPETENCIA Elabora programas en lenguaje Python y bibliotecas a partir del aprendizaje automático y el procesamiento distribuido con Apache Spark. | 3.1 Aprendizaje supervisado con scikit-learn 3.1.1 Clasificación 3.1.2 Regresión 3.2 Aprendizaje no supervisado scikit-learn 3.2.1 Análisis de grupos 3.3 Procesamiento distribuido con Apache Spark 3.3.1 Datos distribuidos resilientes 3.3.2 Transformaciones de datos distribuidos resilientes 3.3.3 Uso de SparkSQL y SparkMLlib | 5.0 | 3.0 | 2.0 |
| | | 5.0 | 3.0 | 2.0 |
| | | 8.0 | 3.0 | 3.0 |
| | Subtotal | 18.0 | 9.0 | 7.0 |

| ESTRATEGIAS DE APRENDIZAJE | EVALUACIÓN DE LOS APRENDIZAJES |
|---|---|
| Aprendizaje Basado en Problemas El estudiante desarrollará las siguientes actividades: 1. Indagación documental del lenguaje de programación Python y sus bibliotecas con lo que elaborarán un mapa conceptual o mental. 2. Se realizarán discusiones dirigidas de lo que obtendrán conclusiones. 3. Soluciona problemas de programación empleando cada tema visto en las unidades temáticas. 4. Elaboración de programas de cómputo que funcionen correctamente utilizando el lenguaje Python y sus bibliotecas. 5. Análisis de casos específicos de los temas vistos 6. Realización de prácticas. | Evaluación diagnóstica Portafolio de evidencias: 1. Mapas mentales/conceptual 2. Conclusión de discusión 3. Problemas resueltos 4. Programas en lenguaje Python 5. Solución de casos 6. Reporte de prácticas |



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Desarrollo de Aplicaciones para análisis de Datos

HOJA 5 DE 7

| RELACIÓN DE PRÁCTICAS | | | |
|------------------------------|---|---------------------------|-----------------------------|
| PRÁCTICA No. | NOMBRE DE LA PRÁCTICA | UNIDADES TEMÁTICAS | LUGAR DE REALIZACIÓN |
| 1 | Aplicación usando ciclos, funciones y archivos con Python | I | Laboratorio de cómputo |
| 2 | Adquisición de datos desde archivos | II | |
| 3 | Diseño de una base de datos con MongoDB | II | |
| 4 | Web scraping con protocolo API-REST | II | |
| 5 | Web scraping accesado por URL usando BeautifulSoup | II | |
| 6 | Web scraping con interacción tipo navegador usando Selenium | II | |
| 7 | Uso de estructuras de datos con Panda | II | |
| 8 | Aprendizaje supervisado con Scikit-learn | III | |
| 9 | Aprendizaje no supervisado con Scikit-learn | III | |
| 10 | Conjunto de datos distribuidos resilientes | III | |
| 11 | Spark SQL | III | |
| 12 | Spark MLlib | III | |
| TOTAL DE HORAS: 27.0 | | | |



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Desarrollo de Aplicaciones para análisis de Datos

HOJA: 6 DE 7

| Bibliografía | | | | | | Documento | |
|---------------------|------------------|------------|--|--------------------------|----------------|------------------|--------------|
| Tipo | Autor(es) | Año | Título del documento | Editorial/ISBN | Librero | Antología | Otros |
| B | Embarak, O. | 2018 | <i>Data Analysis and Visualization Using Python</i> | Apress / 9781484241080 | X | | |
| B | László, G. | 2018 | <i>Website Scraping with Python</i> | Apress / 9781484239247 | X | | |
| B | McKinney, M. | 2018 | <i>Python for Data Analysis</i> | O'Reilly / 9781491957660 | X | | |
| B | Nelly, F. | 2018 | <i>Python Data Analytics</i> | Apress / 9781484239124 | X | | |
| B | Porcu, V. | 2018 | <i>Python for Data Mining Quick Syntax Reference</i> | Apress / 9781484241127 | X | | |

Recursos digitales

| Autor, año, título y Dirección Electrónica | Texto | Simulador | Imagen | Tutorial | Video | Presentación | Diccionario | Otro |
|--|--------------|------------------|---------------|-----------------|--------------|---------------------|--------------------|-------------|
| MongoDB, Inc. (2020). MongoDB. Recuperado el 15 de octubre de 2020, de: https://www.mongodb.com | | | | X | | | | X |
| Python Software Foundation (2020). Python. Recuperado el 15 de octubre de 2020, de: https://www.python.org/downloads/ | | | | X | | | | X |
| The Apache Software Foundation. (2020). Apache Spark Lightning-fast unified analytics engine. Recuperado el 15 de octubre, de: https://spark.apache.org | | | | X | | | | X |



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Desarrollo de Aplicaciones para análisis de Datos

HOJA 7 DE 7

PERFIL DOCENTE: Ingeniero en Sistemas Computacionales, Licenciatura en Computación o áreas afines, con grado de maestría.

| EXPERIENCIA PROFESIONAL | CONOCIMIENTOS | HABILIDADES DIDÁCTICAS | ACTITUDES |
|--|--|---|--|
| Preferentemente 3 años en la industria del software orientado a la Ciencia de Datos y desarrollo de sistemas computacionales. Mínima de 2 años de experiencia en docencia a Nivel Superior. | En desarrollo de sistemas computacionales. Orientados al análisis de datos. Editores de código y entornos de programación orientados a Ciencia de Datos. En herramientas de comunicación y entornos tecnológicos. Del Modelo Educativo Institucional (MEI) | Discursivas Cognoscitivas Metodológicas De conducción del grupo Para evaluar Coordinación del aprendizaje Propicia la investigación Estrategias Metodológicas y Procedimientos | Congruencia Empatía Ética Generosidad Honestidad Proactividad Respeto Responsabilidad Solidaridad Tolerancia Vocación de servicio Compromiso Institucional y social |

ELABORÓ

REVISÓ

AUTORIZÓ

Dr. José Marco Antonio Rueda Meléndez
Profesor Coordinador

M. en C. Andrés Ortigoza Campos
Director ESCOM

M. en C. Mónica Rivera de la Rosa
Profesor colaborador

M. en C. Iván Giovanny Mosso
García
Subdirección Académica
ESCOM

Ing. Carlos Alberto Paredes Treviño
Director UPIIC