

Universidad Nacional Mayor de San Marcos
Universidad del Perú, Decana de América
FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
Centro de Responsabilidad Social y Extensión Universitaria
SÍLABO



PROGRAMA: ESPECIALIZACION EN PYTHON
MÓDULO AVANZADO

I. DATOS GENERALES

Curso : PYTHON MODULO AVANZADO
Total, De Horas : 36 horas (24 de clase + 2 de examen + 10 de trabajos virtuales)
Pre-Requisito : PYTHON INTERMEDIO – MÓDULO II
Costo :

II. SUMILLA

El objetivo fundamental del curso es brindar a los participantes las habilidades y conocimientos esenciales de la ciencia de datos para encontrar patrones en los datos. Utilizando distintas técnicas estadísticas para analizar y extraer información o insights de los datos. A partir de la extracción, el tratamiento y el preprocesamiento de los mismos y que un científico de datos (Data Scientist) debe examinarlos a fondo que se encargará de analizar una gran cantidad de datos para ayudar a las empresas en la toma de decisiones en su estrategia de negocio.

III. OBJETIVOS DEL CURSO

- Al final el curso aprenderás el concepto y uso práctico de la ciencia de datos usando Python y sus librerías conexas;
- Aprenderás a analizar, interpretar datos y crear ventajas competitivas en las organizaciones.
- Aprenderás a comprender la manipulación de datos con Numpy y Pandas,
- Aprenderás a visualizar datos con Matplotlib y Seaborn, reducción de outliers, análisis de agrupamientos, de asociaciones y predicciones, con casos prácticos sobre implementaciones de analítica de datos.
- Aprenderás el flujo para la creación de modelos predictivos para machine learning
- Aprenderás el uso de las librerías necesarias para generar modelos inteligentes.

IV. DIRIGIDO A:

El curso está dirigido a estudiantes universitarios, técnicos y profesionales de Ingeniería de Sistemas, Software, Informática, industrial, TI, Estadística, Economía, administración, matemáticos y profesiones que estén encargados o preparándose implementar, gestionar, utilizar data para la gestión de sus organizaciones.

V. PROGRAMACIÓN:

Introducción
Introducción a la ciencia de datos
Ecosistema de herramientas de DS en Python
Uso de Anaconda y Jupyter
Pandas y Numpy
Uso de Series Object
Práctica: Casos Prácticos

Pre-procesamiento de datos con Numpy y Panda
Introducción a Pandas
Uso de Pandas, Conversión de Series a Dataframe
DataFrame Operations (imputación, normalización, etc)
Análisis de agrupamientos, asociaciones, y reducción dimensional
Limpieza y tratamiento de datos
Carga de datos csv, json, nube.
Práctica: Casos Prácticos

Visualización de datos con Matplotlib y Seaborn
Matplotlib y Seaborn
Estadísticas descriptivas, resúmenes, histogramas y distribuciones
Visualización de datos espaciales y temporales
Tratamiento de archivos CSV
Práctica: Casos Prácticos

Web Scraping
Web Scraping
Introducción de Web Scraping con Python
Ética con web scraping
Obtención de datos en páginas web desplegadas
Práctica: Casos Prácticos

Introducción a Machine Learning
Fuente de datos
Herramientas para aplicar ML en Python
Análisis predictivo
Fundamentos de análisis supervisado y Evaluación de modelos predictivos
Modelos simples de predicción
Práctica: Casos Prácticos

VI. EVALUACIÓN

$$PF = P1 * 0.3 + P2 * 0.3 + EF * 0.4$$

Dónde:

P1 = Práctica 01

P2 = Práctica 02

EF =Examen Final