

Programas Asíncronos

SÍLABO

I. DATOS GENERALES

Curso	:	Preprocesamiento de textos con Python
Área	:	Analytics
Docente	:	Luis Felipe Garayar Burneo
Duración	:	3 horas académicas

II. SUMILLA

El curso se desarrollará mediante la metodología asíncrona, que comprende la revisión de los videos grabados, las lecturas y la resolución de la evaluación planteada. Le permitirá al participante comprender el proceso de extraer información empleando técnicas de análisis de textos como la tokenización, stemming, stopwords, entre otros, previa revisión de los fundamentos y el proceso de preparación de datos. El curso comprende los siguientes temas generales: a) Fundamentos del análisis de textos, b) Procesamiento de textos.

III. REQUISITOS

- Conocimiento de programación en Python.

IV. DIRIGIDO A

- Profesionales, egresados, estudiantes e interesados en comprender el proceso de extraer información empleando técnicas de análisis de textos como la tokenización, stemming, stopwords, entre otros, previa revisión de los fundamentos y el proceso de preparación de datos.

V. OBJETIVOS – LOGROS

El curso le permitirá al participante:

- Comprender los fundamentos del procesamiento de lenguaje natural con Python, así como la tarea de preparación de los datos de cara a este proceso.

- Comprender el proceso de extraer información a partir del análisis de textos a partir de diversas técnicas como la tokenización, stemming, stopwords, entre otros.

VI. METODOLOGÍA

Los contenidos se desarrollarán siguiendo la secuencia teoría – práctica con evaluaciones modulares, utilizando videos grabados y lecturas como recursos didácticos.

VII. CONTENIDOS

Total horas	3 horas
Sesión 1	
• Tema 1: Introducción a Jupyter Notebook y uso de módulo en Python.	20 min.
• Tema 2: Natural Language Processing (NLP): Definición y aplicaciones.	20 min.
• Tema 3: Preparación de datos para procesamiento de textos.	30 min.
Sesión 2	
• Tema 4: Tokenización en Python.	20 min.
• Tema 5: Text cleaning & Case conversions en Python.	30 min.
• Tema 6: Stemming & Lemmatization en Python.	30 min.
• Tema 7: Stopwords en Python.	30 min.

VIII. SISTEMA DE EVALUACIÓN

La nota final del curso se obtendrá a partir de los siguientes rubros de evaluación y pesos asociados:

Rubro	Peso
Evaluación	100%

IX. Docente

Luis Felipe Garayar Burneo

Resumen

- Máster en Inteligencia Artificial – UNIR España
- Magister en Estadística Aplicada – UNALM
- Ingeniero Estadístico – UNI.
- Ingeniero de Sistemas – UNMSM

- Más de 12 años de experiencia en Banca (Inteligencia Comercial: Campañas Dirigidas, Reportería, Análisis y Modelos Estadísticos)