



Análisis Exploratorio de Datos en Python

Para el año 2022 se estima que más del 50% de los negocios globales surjan a partir de los datos (Castellote E. (2019). Es por ello que, para el 88% de los directivos, la inversión en big data e inteligencia artificial es ya una de las mayores prioridades en sus organizaciones (Big Data and AI Executive Survey 2019). Estas cifras nos llevan a reflexionar y entender como en el mundo cada vez se producen y almacenan más y más datos. En la actualidad es imprescindible el uso de la tecnología en todos los procesos desarrollados al interior de las organizaciones.

Gracias a los procesos de transformación digital y tendencias tecnológicas como la industria 4.0, las empresas, además de mantenerse en el mercado, buscan ser cada vez más competitivas y deben integrar a su recurso humano, profesionales con capacidades de análisis y aprovechamiento de datos masivos.

En este aspecto el análisis de los datos es una herramienta muy poderosa que permite a las gerencias de las organizaciones tener acceso a información de valor que se convierte en el insumo más importante para la toma de decisiones.

La prospectiva tecnológica 2016-2025, con base en pronósticos de Gartner e IBM, compara pronósticos de otras reconocidas firmas de analistas como Markess International, Ovum y Pew Research Center, determinando con la hipérbole de Gartner, que la analítica de datos, será el precursor de máquinas inteligentes avanzadas que irán más allá de los protocolos tradicionales de computación capaces de percibir y aprender por sí solas [FEDESOFTE 2015].

En este sentido, el SENA ofrece el programa de formación complementaria en análisis exploratorio de datos, con el objetivo de brindar una formación que permita adquirir habilidades y destrezas en técnicas de análisis y exploración de datos utilizando para este proceso el lenguaje de programación python, de tal manera que los aprendices en esta área contribuyan al sector productivo encontrado ventajas en el aprovechamiento de sus datos para la toma de decisiones y la predicción de comportamientos futuros, basados en análisis estadístico y la aplicación de técnicas propias del campo de la analítica de datos.



Contenido

RA1.

- Entender el negocio: comprender la misión, visión y metas empresariales, establecimiento de objetivos a corto, mediano y largo plazo, proyección y análisis de expansión y crecimiento futuro.
- Datos: fuentes de datos e información, exploración de la estructura de los datos.
- Preparar la data: preparación de datos, detectar errores y datos faltantes, identificación de las variables más importantes, definir las preguntas básicas, verificar si los valores de las variables están en la escala correcta, entender la data.
- Introducción al análisis exploratorio de datos: estadística descriptiva y estadística inferencial, población y muestra, carácter, variable y atributo, la medida, escalas de medida y clasificación de variables.

RA2.

- Métodos para hacer análisis exploratorio de datos: datos univariados, datos bivariados, datos multivariados.
- Estructuras de datos: series y dataframes, dataset, ordenamiento de datos, indexación y recuperación.
- Herramientas tecnológicas para el análisis y visualización de datos: python, librerías (pandas, numpy, matplotlib, sklearn, entre otras), ide google collaborate, jupyter, spyder, pycharm, anaconda, orange.
- Visualización básica de datos: gráficos de barra, gráficos de columna, histogramas.

A3.

- Evaluación de resultados: evaluación del resultado, comprobar el resultado obtenido frente a los objetivos del negocio.
- Validación preguntas de análisis: comprobación del resultado del modelo frente a las preguntas de análisis propuestas, aceptación o propuesta de nuevos interrogantes.



Habilidades que Desarrolla

RA1:

- Recopilar información acerca de la situación actual de la empresa o negocio objeto de estudio.
- Identificar los objetivos de negocio de la empresa u organización objeto de análisis.
- Identificar fuentes de datos.
- Preparar los datos para su análisis en busca de errores y datos faltantes.
- Definir las preguntas objeto de análisis.

RA2:

- Definir el método para hacer el análisis exploratorio de los datos.
- Preparar los datos recopilados de diferentes fuentes ordenando e indexando los datos.
- Preparar el dataset o base de datos a ser analizada.
- Determinar las variables objeto de análisis.
- Generar el modelo de análisis.
- Preparar la infraestructura tecnológica a ser utilizada en el modelo propuesto.
- Definir el tipo de grafico que permita visualizar las variables analizadas.
- Generar grafico que permita visualizar resultados del análisis.

Ra3:

- Evaluar el modelo construido.
- Comprobar el resultado del modelo vs los objetivos de la organización modelo de mejora continua.
- Verificar si el resultado del modelo de análisis da respuesta a las preguntas de análisis planteadas para la empresa u organización.
- Preparar informe del análisis exploratorio de los datos.



Requisitos de Ingreso

Conocimientos en:

- Preferiblemente conocimientos en lógica de programación y algoritmia básica (estructura general de un algoritmo, variables, constantes, estructuras condicionales, estructuras repetitivas).
- Recomendable tener conocimientos básicos en estadística como: Población, muestra, individuo, muestreo, valor, dato, máximo, mínimo, promedio, varianza.
- Recomendable haber realizado previamente el programa Variables y Estructuras de Control en Python.
- Competencias en solución de problemas con algoritmia y lógica de programación.

Académicos:

Grado 9 aprobado,

Requisito adicional: Superar prueba de aptitud, motivación, interés y competencias mínimas de ingreso.

Requisito de ley: Edad mínima definida en la ley (14 años).

Inscríbese aquí! 