

CIENCIA DE DATOS

ANALISIS COMPARATIVO CIENCIA DE DATOS

EDISON ANDRÉS FORERO RIAÑO

DAMIAN DANILO NARANJO PERILLA

ESCUELA TECNOLÓGICA INSTITUTO TÉCNICO CENTRAL

FACULTAD DE INGENIERÍA

TECNOLOGÍA EN DESARROLLO DE SOFTWARE

BOGOTÁ D.C

2024

La visualización de datos en ciencia de datos y big data utiliza diferentes métodos para adaptarse al volumen de datos y a las necesidades de investigación específicas. En ciencia de datos, trabajamos con datos estructurados y no estructurados, lo que permite un análisis detallado y eficiente. Un ejemplo bien conocido es el conjunto de datos del Titanic, que contiene variables fácilmente manipulables como clase, edad y sexo. Para este tipo de datos, la visualización se centra en proporcionar información clara y específica, como tasas de supervivencia entre diferentes categorías de viajeros. Los cuadros, histogramas, gráficos de barras, etc. ayudan a resaltar diferentes patrones y tendencias. Esta capacidad de profundizar en los datos permite a los científicos de datos realizar análisis detallados que proporcionan información clara y específica para la toma de decisiones informadas.

En contraste, en el contexto de Big Data, se manejan volúmenes masivos y a menudo no estructurados de datos, lo que exige un enfoque visual más generalizado. Aquí, la visualización se utiliza para capturar patrones amplios en grandes flujos de información, como los generados por las redes sociales, los sensores IoT o las plataformas de comercio electrónico. Dado que el procesamiento de datos en Big Data es continuo y a gran escala, la visualización debe optimizarse para proporcionar resúmenes y tendencias en tiempo real. Se utilizan algoritmos avanzados y técnicas de automatización para generar visualizaciones que destacan de manera rápida y eficiente los patrones de comportamiento a gran escala. Este es un factor clave para facilitar la toma de decisiones cuando hay demasiada información para analizar cada parte en detalle.

Si bien el análisis de la ciencia de datos permite a los investigadores analizar datos para identificar relaciones y patrones detallados, el objetivo del big data es describir rápidamente patrones generales que respaldan las decisiones estratégicas. En resumen, si bien ambos campos utilizan la visualización de datos, la ciencia de datos se centra en el análisis detallado y profundo de la información, mientras que los macrodatos buscan identificar patrones amplios y proporcionar explicaciones de las condiciones, el tiempo y la escala cambiantes.