

Hadoop

Es un marco de trabajo de código abierto diseñado para almacenar y procesar grandes volúmenes de datos en clusters de servidores. Utiliza el sistema de archivos distribuido para guardar información de manera replicada y confiable.

Clústeres

Es un conjunto de computadoras interconectadas que trabajan como si fueran una sola. En el contexto de Big Data, frameworks como Hadoop y Spark aprovechan los clústeres para dividir y procesar grandes volúmenes de datos en paralelo.

Big Data

Se refiere al manejo y análisis de grandes volúmenes de datos que son demasiado extensos para ser procesados por sistemas tradicionales. Este concepto no solo abarca la cantidad de información, sino también la velocidad con la que se genera.

BIG DATA Y ANALISIS DE DATOS

Spark

Poderosa plataforma de procesamiento de datos diseñada para trabajar con grandes volúmenes de información. Spark realiza la mayoría de las operaciones en memoria, lo que le permite ser mucho más veloz en tareas de análisis.

Mercado

Se utiliza en diversos entornos como la salud, para analizar historiales clínicos y predecir enfermedades; en los bancos, para detectar fraudes y evaluar riesgos; en el comercio electrónico, para recomendar productos y personalizar ofertas.

Seguridad

Se utilizan técnicas como cifrado, autenticación, control de accesos, auditorías y monitoreo constante, la seguridad en Big Data no solo protege la información, sino que también garantiza la confianza y el cumplimiento de normativas legales.