Basic Data Concepts:

Introduction to Data and Data Processing

Objectives

Objective

Explain data models and data processing concepts.

What is Data?

1.79 billion active Facebook users

4.6 billion mobile phone users in 2016

2.8 trillion sensor devices by 2019

50% of city residents will interest with Smart City by 2019

Smart home market $43 billion by 2020

7 million drones in the US Skies by 2020

10 million self-driving cars will be on the road by 2020

200 million wearable devices will be shipped by 2020

What is data processing?

Data → Process → Information

What is big data?

- The volume of available data is huge

- The data keeps growing at a staggering rate

- Data comes from a variety of sources in totally different formats

BIG DATA

What is scalable data processing?

Allows database processing systems to cope with the volume, velocity and variety aspects that big data brings into the system.

A plethora of data processing systems

Relational DBMS(Centralized, Distributed)

– Operational workload (bank, online store)

NoSQL graph, document, key-value stores)

– Unstructured data (highly available systems)

Hadoop, Spark

– For analytics over massive-scale data