Presented By:

Anish Poudel

Karun Sunuwar

Sabin Pandey

Sarad Adhikari

Introduction

- It refers to vast and complex datasets.
- Combination of structured, semi-structured, and unstructured data.
- It is characterized by 5 V’s (Volume, Velocity, Variety, Veracity, Value).
- Social Media, Machine and Transactional Data, Research, Web logs.
- Can exceed up to terabytes or petabytes.

Core Technology

Distributed Computing Frameworks:

- Apache Hadoop
- Apache Spark

Data Storage Technologies:

- Hadoop Distributed File System (HDFS)
- Apache Cassandra

Data Analytics Tools:

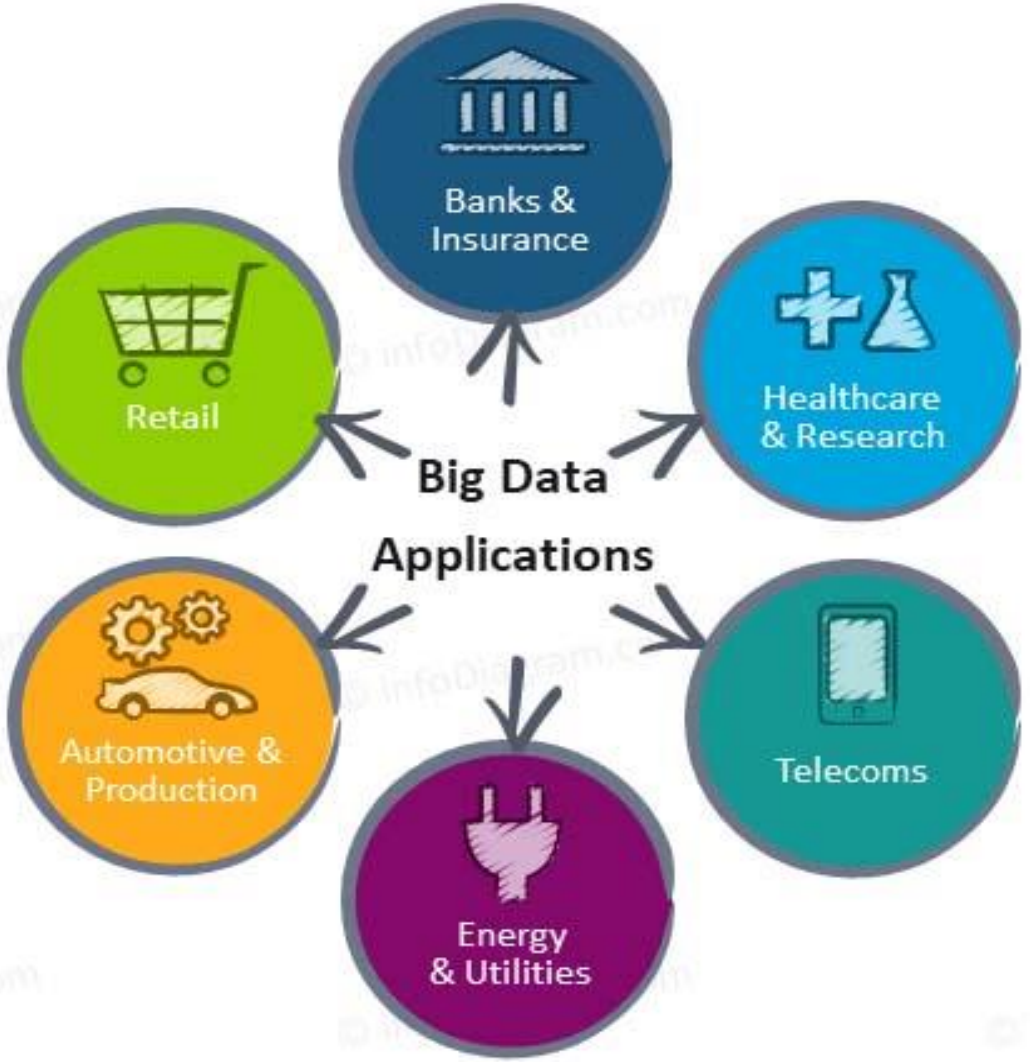
- Machine Learning Libraries (TensorFlow, PyTorch)
- Business Intelligence Tools

Data Visualization Tools:

- Tableau, Power BI and other platforms

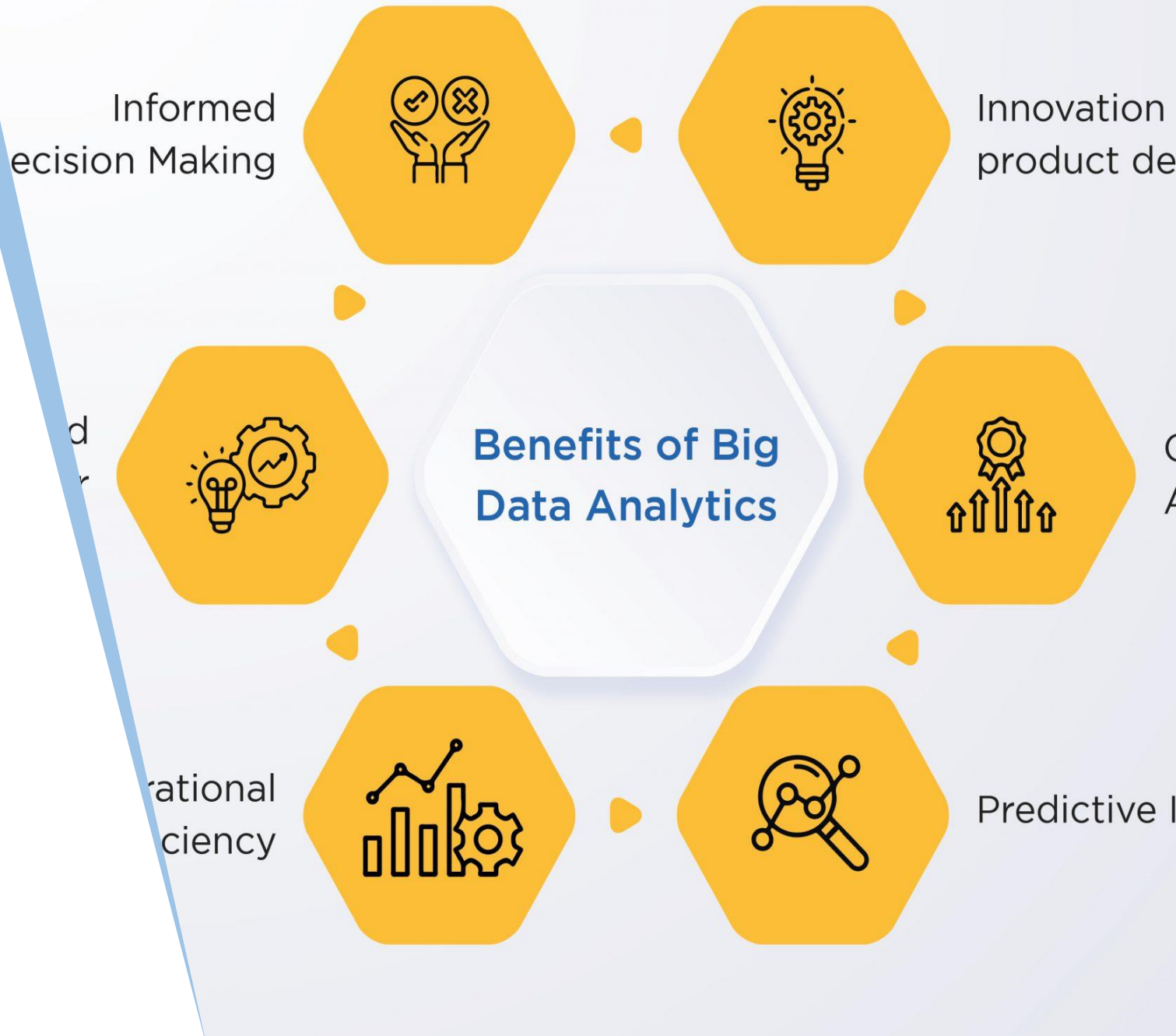
Application

- Recommendation System
- Fraud Detection
- Healthcare
- Traffic Predictions



Benefits

- Improved decision-making
- Improved risk management
- More efficient operations
- Better customer experiences
- Increased agility and innovation



Challenges

- Security concerns.
- Speed of data growth.
- Lack of data talent and skills.
- Problems with data quality.

Additional Information

In every seconds Google (43000 above search) are seen , Youtube(83000 views)Email sent(2329749 and 21+tb of Internet TrafficAll the online platforms social media platforms(Facebook, Netflix ,amazon, Twitter etc.) uses big data to perform meaningful data analytic, Example of google map it has a lot of information about the every single location of the entire worldAlso generates the new data advantages:-Tracking consumer, Monitoring payment patterns Using image data from cameras and sensors, as well as GPS data, to detect problems and improve road maintenance in cities Working Integration ManagementAnalysis