

SNS COLLEGE OF TECHNOLOGY

**An Autonomous Institution
Coimbatore-35**



DEPARTMENT OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE

23ADT202 – FUNDAMENTALS OF DATA SCIENCE AND ANALYTICS

II YEAR IV SEM

UNIT I – INTRODUCTION TO DATA SCIENCE

BENRFITS AND USES OF DATA

Benefits and Uses of Data Science

Empathy:

- Improved Decision Making

Data science enables organizations to move from intuition-based decisions to evidence-based decisions by analyzing historical and real-time data.

- Extraction of Hidden Knowledge

Large datasets often contain patterns that are not visible through traditional analysis. Data science helps discover hidden relationships, trends, and anomalies.

- Predictive Capability

Using statistical models and machine learning, data science allows organizations to predict future events, such as customer behavior, system failures, or market trends.

- Handling Big and Complex Data

The book emphasises that modern data is large, fast, and diverse. Data science offers methods for efficiently processing both structured and unstructured data.

- Automation of Analytical Tasks

Repetitive analysis and complex decision processes can be automated using data-driven models, increasing speed and consistency.

- Competitive Advantage

Organisations that effectively utilise data science gain a strategic advantage by optimising operations, enhancing products, and gaining a deeper understanding of their customers.

Problem Statement:

Define:

- Organizations across industries generate massive volumes of structured and unstructured data from transactions, sensors, social media, and digital platforms.
- However, much of this data remains not effectively used due to the lack of effective analytical approaches.
- Decision-making is often based on intuition or limited reports rather than data-driven insights, leading to inefficiencies, increased costs, poor customer satisfaction, and missed opportunities.
- There is a need to systematically use Data Science techniques to extract meaningful insights, predict outcomes, optimise operations, and support strategic decision-making.

Uses of Data Science

- Business Intelligence and Analytics

Data science supports reporting, dashboards, and advanced analytics to monitor performance and identify opportunities for improvement.

- Customer Behavior Analysis

By analyzing transaction data, logs, and interactions, data science helps understand customer preferences, churn, and satisfaction.

- Risk and Fraud Detection

Statistical and machine learning models identify unusual patterns, helping detect fraud, credit risk, and operational risks.

- Recommendation Systems

Data science techniques are used to recommend products, services, or content based on user behavior and historical data.

Ideate:

- Apply Data Science techniques such as **data mining, machine learning, and predictive analytics**.
- Use visualization tools to present insights in an understandable form.
- Develop models for forecasting demand, detecting fraud, and predicting failures.
- Implement personalization systems using customer data. Integrate Data Science solutions into existing business processes.

Prototype:

- Build a data pipeline for data collection, cleaning, and preprocessing.
- Develop predictive models (e.g., sales forecasting, risk prediction).

- Operational Optimization

Industries use data science to optimize supply chains, resource allocation, scheduling, and maintenance.

- Scientific and Research Applications

The book highlights data science as a key driver in scientific discovery, enabling researchers to analyze experimental and observational data at scale.

- Predictive Maintenance

Sensor and machine data are analyzed to predict failures before they occur, reducing downtime and costs.

Design Thinking–Based Solution

- Create dashboards to visualize trends and patterns.
- Implement recommendation or classification systems. Test the system on historical and real-time datasets.

Testing:

- Evaluate model performance using accuracy, precision, and recall. Validate predictions against real-world outcomes. Gather feedback from stakeholders and end users.
- Improve models based on errors and new data. Measure impact in terms of cost reduction, efficiency, and decision quality.