Use Cases of Data Science

1. E-commerce

Data Source: Customer interactions, transaction data, product reviews, social media feedback

Data Issues: Incomplete profiles, sparse data for new users, noisy or biased reviews

Type of Data: Structured (transactions), unstructured (reviews, social media posts)

Problem Statement: How can we build personalized recommendations to improve user experience and boost sales?

Solution: Recommender systems using collaborative filtering, content-based models, and deep learning.

2. Manufacturing

Data Source: IoT sensors, production logs, quality control data, maintenance records

Data Issues: Sensor failures, high dimensionality, real-time data processing challenges

Type of Data: Time-series, semi-structured, structured

Problem Statement: How can we predict equipment failures and reduce unexpected downtime?

Solution: Predictive maintenance models leveraging ML algorithms like Random Forest and LSTM.

3. Banking

Data Source: Transaction records, customer profiles, credit scores, external economic indicators

Data Issues: Imbalanced fraud data, privacy concerns, data drift over time

Type of Data: Structured, tabular, time-series

Problem Statement: How can banks detect fraudulent transactions in real time?

Solution: Fraud detection systems using anomaly detection, logistic regression, and ensemble methods.

4. Healthcare

Data Source: EHRs, medical imaging data, wearable health devices

Data Issues: Missing/inconsistent records, privacy concerns, bias in clinical trial data

Type of Data: Structured (patient history), unstructured (doctor’s notes, images)

Problem Statement: How can we use patient data to predict disease risk and recommend early interventions?

Solution: Disease prediction models using deep learning and NLP.

5. Transport

Data Source: GPS data, traffic sensors, ride-hailing app data

Data Issues: High velocity and volume, GPS inaccuracies, unpredictable external factors

Type of Data: Geospatial, time-series

Problem Statement: How can transport companies optimize routes and predict traffic patterns?

Solution: Traffic prediction models using ARIMA, LSTMs, and graph neural networks.

6. Finance

Data Source: Stock market data, financial statements, news articles, social media

Data Issues: Market noise, sudden events, unstructured text from news sources

Type of Data: Time-series, structured, unstructured (text data)

Problem Statement: How can we build a trading strategy that maximizes returns and minimizes risk?

Solution: Stock price forecasting models with deep learning and sentiment analysis.