Unit 4: Data Refinement and Transformation

1. Explain the importance of data quality in data science projects with examples.
2. Discuss the role of data profiling in understanding a dataset before analysis.
3. Analyze the challenges faced when dealing with incomplete data and suggest techniques to manage them.
4. Explain the concept of missing data imputation and provide real-world examples where it is necessary.
5. Discuss how Moving Average Smoothing can be applied to remove noise from a dataset.
6. Analyze the role of Gaussian Filtering in smoothing data for pattern recognition tasks.
7. Explain the process of categorization in datasets and how it enhances data quality.
8. Discuss the working of One Hot Encoding and its relevance in machine learning models with an example.
9. Analyze the differences between label encoding and one hot encoding, and suggest where each should be applied.
10. Explain how poor data refinement affects the performance of machine learning models with real-time applications like fraud detection.

Unit 5: Introduction to Time Series and Text Data Analysis

1. Explain the key characteristics of time-dependent data and why they differ from regular datasets.
2. Discuss the significance of timestamps in time series analysis with real-world examples like stock market analysis.
3. Analyze how seasonal patterns can be identified in time series datasets and give a practical example.
4. Discuss the importance of autocorrelation plots in understanding time series data behavior.
5. Analyze the impact of missing time intervals in time series data and strategies to handle them.
6. Explain the steps involved in trend analysis and how it supports forecasting future events.
7. Discuss an example where trend analysis helped a business make a strategic decision (e.g., retail sales forecasting).
8. Analyze how decomposition of time series data into trend, seasonality, and residual helps in better model building.
9. Explain the applications of time series analysis in sectors like healthcare, energy, and finance.

Unit 6: Advanced Concepts and Applications in Data Science

1. Explain the steps involved in executing a real-world data science project from start to end.
2. Discuss how machine learning is used in recommendation systems with examples like Netflix or Amazon.
3. Explain the importance of exploratory data analysis (EDA) in preparing a dataset for a mini project.
4. Discuss how a mini project using Python can help strengthen practical knowledge of data science concepts.
5. Analyze the role of data preprocessing in small dataset analysis and suggest suitable Python libraries.
6. Explain ethical challenges faced in data science like bias, fairness, and privacy issues with real-time examples.
7. Discuss the importance of data anonymization and how it is applied in industries like healthcare and banking.

8. Analyze a real-world scenario where unethical use of data led to a major scandal (example: Cambridge Analytica case).
9. Explain the significance of transparent and explainable AI models in building trust among users