Data Science Important Question

Basic Concepts

- What is Data Science? Explain its life cycle.
- Difference between Data Science and Data Analytics.
- What are the various types of data (structured, semi-structured, unstructured)?
- What is data cleaning and why is it important?
- Explain feature engineering and feature selection.
- What are KPI, lift, model fitting, robustness, and DOE?
- Explain the different learning types in machine learning (supervised, unsupervised, reinforcement).

Frequently Asked Interview Questions

- What is overfitting and how can you prevent it?
- What is a confusion matrix? Explain precision, recall, and F1 score.
- Describe the role of activation functions in machine learning models.
- Explain regularization techniques and why they are necessary.
- What is cross-validation and why is it used?
- Difference between bagging and boosting algorithms.
- What is dimensionality reduction? Explain PCA and t-SNE.
- What are classification and regression problems?
- Explain cluster analysis and types of clustering algorithms.

University Exam Questions

- Write and explain a Python program for data manipulation using libraries like Pandas.
- Explain descriptive and inferential statistics with examples.
- Describe supervised learning algorithms, including decision trees and support vector machines.
- Explain the concept of hypothesis testing and p-value.
- Discuss exploratory data analysis and its role.
- Write note on network analysis and social network analysis.
- Describe regression analysis and its types.
- Explain centrality and community detection in networks.

Long and Conceptual Questions

- Discuss the importance of data preparation and cleaning.
- Explain the machine learning pipeline from data gathering to model deployment.
- Describe different evaluation metrics for classification problems.
- Explain the concept of neural networks and deep learning.
- Discuss ethical considerations in data science.
- Explain time series analysis and forecasting.
- Describe the working and applications of recommender systems.

Coding and Practical Questions

- How to read and write CSV, Excel files in Python?
- Implement and explain confusion matrix calculations.
- Write a program for decision tree classification.
- Explain random forest algorithm functionality and use cases.
- Implement clustering using K-means in Python.