

1. Discuss the process and significance of cross-validation for time series (e.g., rolling window). Why is standard k-fold not valid?
2. What is normalization? Describe Min–Max Normalization with an example.
3. Compare Bagging and Boosting with examples of algorithms using them.
4. Describe the working principles of Fourier Transform, Wavelet Transform, and EMD for time series decomposition. Compare their use-cases.
5. Explain any two methods of handling missing values in a dataset.
6. Describe the assumptions of Linear Regression and explain why they matter.
7. Explain Hybrid Forecasting Models (Classical + ML). Give an example of an additive and a multiplicative hybrid approach.
8. Define Trend and Seasonality in time series with simple examples.
9. Explain the concept of Gradient Boosting. How does XGBoost improve traditional boosting?
10. A dataset exhibits complex seasonality, trend shifts, and noise. Propose a complete forecasting solution mixing decomposition, ML, and statistical models, and justify each step.
11. What is a Random Forest? How does it reduce overfitting?
12. Discuss the steps involved in SARIMA model construction and when it is preferred.
13. Explain Gini Index and Information Gain used for attribute selection in decision trees.
14. What is Logistic Regression? Explain its use in binary classification.
15. Explain how Trend, Seasonality, and Cyclic components can be visually analyzed in time series data.
16. Discuss the challenges of applying Random Forests and Boosting on temporal data, and propose solutions such as sliding windows and feature engineering.
17. What is a Decision Tree? Mention any two advantages of using it.
18. Describe the model selection criteria AIC & BIC and their role in ARIMA modeling.

19. Explain with diagrams how additive and multiplicative decomposition differ and how they impact forecasting strategies.
20. Write short notes on: a) Noise component in time series b) Differencing
21. Explain in detail the steps of outlier detection and treatment with suitable examples.
22. Discuss in detail how machine learning models (LR, RF, XGBoost) can be adapted for time series forecasting. Include advantages and limitations.
23. Define RMSE, MAE, and R² used in regression model evaluation.
24. Explain the process of identifying parameters (p, d, q) for an ARIMA model.
25. Compare ARIMA/SARIMA with Machine Learning models in terms of assumptions, nonlinearity handling, interpretability, and compute cost.
26. What do you mean by Stationarity in time series? Why is it important?
27. Write a detailed note on hyperparameter tuning and its importance in ensemble models.
28. Describe the role of the confusion matrix in model evaluation.
29. Design a detailed model evaluation pipeline combining feature importance analysis, residual diagnostics, and forecast error metrics.
30. Explain the process and significance of cross-validation for time series (already listed) – wait, duplicate; instead use: Explain the concept of Gradient Boosting – duplicate too.