

Group Project Submission Guidelines: Time Series Analysis

This document outlines the submission guidelines for the Time Series Analysis Group Project. Teams of 5-6 students need to be formed for the project development and project should be developing using Python.

Project Goal:

The goal of this project is to apply time series analysis techniques to a chosen dataset and build models for forecasting.

Deliverables:

1. **PowerPoint Presentation (12+3 Minutes):** Live presentation during the last class.
2. **Python Code (on GitHub Classroom):** Complete code for all project stages.

Project Stages:

The project will involve the following stages:

1. Data Acquisition (10%):

- Choose a publicly available dataset relevant to time series analysis.
- Briefly explain why you selected this dataset and what is the problem statement.
- Document the data source, including a brief description and link to the data source.

2. Data Preprocessing (10%):

- Describe the steps taken for data cleaning, including handling missing values, outliers, and inconsistencies.
- Explain any transformations performed on the data (e.g., scaling, differencing).
- Visualize the data before and after preprocessing to demonstrate the impact of cleaning steps.

3. Exploratory Data Analysis (20%):

- Perform basic exploratory data analysis (EDA) to understand the data's characteristics.
- Analyze time series properties like stationarity, trend, and seasonality using relevant visualizations.
- Calculate and interpret descriptive statistics relevant to the chosen analysis type.

4. Time Series Modeling (30%):

- Implement univariate time series forecasting models like Benchmark models, Exponential Smoothing methods, Theta, ARIMA family of models.
- Consider multivariate models such as Prophet, (S)ARIMAX and Vector Autoregression methods if your data has relevant features that might influence the target variable.

- Explain the model selection process and justify your choice of models.
- Train and evaluate the models using appropriate performance metrics (e.g., RMSE, MAE, MAPE).

5. Time Series Cross-Validation (10%):

- Implement the time series cross-validation technique to assess the model's generalizability.
- Explain the chosen cross-validation method and its importance in time series analysis.
- Report the cross-validated performance metrics for a more robust evaluation.

6. Residual Diagnostics (10%):

- Analyze the model's residuals to identify any potential issues like autocorrelations or heteroscedasticity.
- Apply appropriate tests and visualizations to assess the model's assumptions.
- Discuss any deviations from assumptions and potential consequences for model accuracy.

7. Error Reporting and Discussion (10%):

- Report the final performance metrics achieved through the chosen models.
- Discuss the limitations of the chosen dataset and models.
- Suggest potential areas for improvement or further analysis with additional data or techniques.

PowerPoint Presentation:

- The presentation should be clear, concise, and visually appealing (15 slides max.).
- Introduce the project topic, data, and preprocessing steps.
- Explain the chosen time series models and their implementation.
- Present the results, including visualizations of model performance and residual analysis.
- Discuss the limitations, conclusions, and future directions.
- Allocate time for questions and discussion.

Python Code on GitHub Classroom:

- Push a well-organized and documented Python code repository to GitHub Classroom.
- The code should cover all project stages, including data loading, preprocessing, model building, evaluation, and visualization.
- Use clear variable names, comments, and functions for better readability.
- Ensure the code is well-structured and easy to follow for others.

Grading Rubric:

- Project Work (as detailed in Project Stages section) (80%)
- Presentation: Clarity, structure, visuals, and Q&A (10%)

- Code Quality: Organization, documentation, and functionality (10%)

Additional Notes:

- All group members are expected to contribute to the project and be prepared to answer questions during the presentation.
- Ensure proper citation of any external resources used in the project.
- Maintain proper academic integrity throughout the project.

The Group Project must be submitted by **21st June '24 (Friday) 9am** before start of class.
