Semester Project Final Deliverable

CS 637 Time Series and Forecasting, Fall 2021

Our final project deliverable will consist of two parts; A written paper worth of findings and discoveries through your work (which would also include your code as a .RMD file) and a presentation to display your work for the class. The expectations of both parts are outlined below

Written Paper & Code

The written paper should be a synthesis of your finding from the three deliverables. In particular, we are looking to have each point addressed below:

1. One paragraph on defining your problem
    a. Which domain does this apply to?
    b. What are you trying to accomplish?
    c. Why are you trying to accomplish this goal?
2. One paragraph on defining the dataset
    a. Where was the data accessed from?
    b. How many observations and variables does the data include?
    c. How does the data agree with you solving your problem statement?
3. Discussion of correlation within the data
    a. Correlogram
    b. Variance
    c. Auto Correlation
    d. Appropriate visuals for each topic
4. A discussion on forecasting strategies covered in Chapter 3
5. Implementation or discussion on a basic stochastic model
    a. Determine if auto-regressive model is appropriate
        i. What is the number of lag terms?
    b. Determine alpha coefficient
    c. Build ACF plot

6. Build a Linear Regression

    d. We are looking for you to implement one (if not multiple) of the types of regression models we discussed in class. Include results and visualizations
        i. Generalized least squares
        ii. Linear models with season variables
        iii. Harmonic seasonal models
        iv. Log Transform regression
        v. Non linear model
7. Implement or discuss a moving average model, seasonal or non-seasonal ARIMA model, or GARCH model
    a. Any of these models is fine to choose. Discuss implementation and results from the model of your choosing, including appropriate visuals

8. Include one .RMD file with all code, commented as appropriate, with code that transformed data, computed models, and built visualizations

Presentation

We will slot time for each individual to present their work, for a total of 5-7 minutes. We will reserve 3-4 minutes for Q&A afterward, as well as hearing feedback from your classmates. Your presentation can be of a format of your choosing, with PowerPoint being the most accessible and popular option. We are looking for you to cover a few objectives as you present to your classmates

1. Introduction of the problem
    a. Motivation to choose this problem
    b. Why it is important to be modeled
2. Structure of your data
    a. How did you find/access your data?
    b. What type of manipulation did you have to do?
    c. Maybe include summary of data statistics (distributions, mean, variance, etc.)
3. Correlation in the data
    a. Include visuals on correlation and auto-correlation and what it looked like in your data
4. Model building
    a. Discuss each model implemented
        i. Why was this model chosen
        ii. Did any of the data required to be transformed/manipulated before using the model
    b. Visualizations from each model
    c. Results from each model
5. Please include all other relevant work you put into your project that you believe would be beneficial to share with the class