ADS 506 Module 2 Exercises: Chapter 4

This assignment is due on Day 7 of the learning week. The assignment for this module consists of written work. Complete this entire assignment in R Markdown. You will need to include the question and number that you are answering within your submitted assignment. **Once completed, you will knit your deliverable to a Word/PDF file.**

Chapter 4: Forecasting Methods: Overview (Pages 77-78) #1 and 2

1. A large medical clinic would like to forecast daily patient visits for purposes of staffing.

    a. If data are available only for the last month, how does this affect the choice of model-based vs. data-driven methods?
    b. The clinic has access to the admissions data of a nearby hospital. Under what conditions will including the hospital information be potentially useful for forecasting the clinic's daily visits?
    c. Thus far, the clinic administrator takes a heuristic approach, using the visit numbers from the same day of the previous week as a forecast. What is the advantage of this approach? What is the disadvantage?
    d. What level of automation appears to be required for this task? Explain.
    e. Describe two approaches for improving the current heuristic (naive) forecasting approach using ensembles.

2. The ability to scale up renewable energy, and in particular wind power and speed, is dependent on the ability to forecast its short-term availability. Soman et al. (2010) describe different methods for wind power forecasting (the quote is slightly edited for brevity):

    Persistence Method: This method is also known as "Naive Predictor". It is assumed that the wind speed at time $t + \delta t$ will be the same as it was at time t. Unbelievably, it is more accurate than most of the physical and statistical methods for very-short to short term forecasts.

    Physical Approach: Physical systems use parameterizations based on a detailed physical description of the atmosphere.

    Statistical Approach: The statistical approach is based on training with measurement data and uses difference between the predicted and the actual wind speeds in the

immediate past to tune model parameters. It is easy to model, inexpensive, and provides timely predictions. It is not based on any predefined mathematical model and rather it is based on patterns.
Hybrid Approach: In general, the combination of different approaches such as mixing physical and statistical approaches or combining short term and medium term models, etc., is referred to as the hybrid approach.

a. For each of the four types of methods, describe whether it is model-based, data-driven, or a combination .
b. For each of the four types of methods, describe whether it is based on extrapolation, causal modeling, correlation modeling, or a combination.
c. Describe the advantages and disadvantages of the hybrid approach.