Homework 5

Overview

In this assignment, you will reflect on one of the readings from Week 8 on either model interpretability or model fairness. **This assignment is to be completed individually.**

Objectives

- Gain experience reading and interpretting peer-reviewed academic publications
- Describe, at a high level, the findings presented in an academic paper
- Connect a paper's findings back to your own data mining work

Grading

- Uploaded requested files, 5%
- File is properly/clearly formatted, 5%
- Writing is clear and appropriately formal, 5%
- Part A
    - Addresses Q1: 5%
    - Addresses Q2: 10%
    - Addresses Q2: 35%
    - Addresses Q3: 35%

Deliverables

- A .pdf file with your responses

Formatting

Include headings for each question with each response underneath the appropriate heading. You do not need to use R markdown to generate your pdf for this assignment.

Part A.

For Part A, choose one of the readings from Week 8 on either model interpretability or model fairness that you found most interesting or useful (one of the articles listed below). Each of these articles is uploaded as a pdf under Week 8's weekly content.

1. Burrell, Jenna. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." Big Data & Society 3, no. 1 (June 1, 2016): 205395171562251. https://doi.org/10.1177/2053951715622512.
2. Poursabzi-Sangdeh, Forough, Daniel G Goldstein, Jake M Hofman, Jennifer Wortman Wortman Vaughan, and Hanna Wallach. "Manipulating and Measuring Model Interpretability." In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 1–52. Yokohama Japan: ACM, 2021. https://doi.org/10.1145/3411764.3445315.
3. Goodman, Bryce, and Seth Flaxman. "European Union Regulations on Algorithmic Decision-Making and a 'Right to Explanation.'" AI Magazine 38, no. 3 (October 2, 2017): 50–57. https://doi.org/10.1609/aimag.v38i3.2741.

4. Birhane, Abeba. "Algorithmic Injustice: A Relational Ethics Approach." Patterns 2, no. 2 (February 2021): 100205. https://doi.org/10.1016/j.patter.2021.100205.
5. Mitchell, Margaret, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. "Model Cards for Model Reporting." In Proceedings of the Conference on Fairness, Accountability, and Transparency, 220–29. Atlanta GA USA: ACM, 2019. https://doi.org/10.1145/3287560.3287596.

Answer the following questions. Your responses should be in complete sentences. None of your responses need to be more than a paragraph, but it should be clear that you have read and made an effort to understand the chosen paper.

1. Which paper did you select?
2. Why did you choose to read the paper you selected?
3. In your own words (1 - 2 paragraphs), what is the paper about? E.g., what are the goals and/or findings?
4. How could you apply the ideas presented in the article to your own future data mining projects (in this class and/or beyond the classroom)? Provide specific examples.