

Part 3: Critical Thinking (20 points)

Ethics and Bias (10 points)

How biased training data may affect patient outcomes in the case study

Biased data may result in inappropriately skewed predictions that incorrectly classify targeted groups of patients. For example, say the training set does not contain many women or rural patients, so every time it will underestimate their likelihood of being readmitted. This may lead to poor patients who lack follow-up care, resulting in further health disparities.

There is one strategy to mitigate this bias;

Use insightful data augmentation and model testing. This includes:

- i. Conducting the subgroup analysis to observe balanced performance among demographics.
- ii. Increasing the representation of the samples in the dataset.
- iii. Adding fairness metrics (e.g., equal opportunity difference) to the evaluation process.

Trade-offs (10 points)

Trade-off between model interpretability and accuracy in healthcare.

When dealing with healthcare, the interpretability of a model is usually preferred over accuracy since clinicians must be able to comprehend and trust the model's actions. An extremely precise black-box (such as a deep neural network) model might not suffice when it cannot explain its computation. Such models as XGBoost find a nice golden mean, providing relatively good results and insights into the importance of the features.

Influence of scarce computational resources on model choice:

In case the hospital does not have robust resources (no GPUs, bare servers only), server-heavy models such as logistic regression or a decision tree are better to consider than compute-intensive models. Such uncomplicated models are also more quickly trained, easily deployed in local systems, and have less dependence, which can be handy in the real-world hospital settings.