**📘 AI for Software Engineering**

**🧠 Part 3: Critical Thinking (20 Points)**

**🔍 Ethics & Bias (10 points)**

**Q: How might biased training data affect patient outcomes in the case study?**  
Biased training data can result in unfair treatment recommendations. For instance, if the dataset over represents a specific demographic (e.g., older males), the AI may under predict readmission risks for underrepresented groups such as women or ethnic minorities. This could lead to missed follow-up care or misallocation of medical resources, ultimately compromising patient safety and worsening health disparities.

**Q: Suggest one strategy to mitigate this bias.**  
One effective strategy is to apply **re-sampling techniques**, such as **SMOTE (Synthetic Minority Oversampling Technique)**, to balance the dataset. Additionally, tools like **IBM AI Fairness 360** can evaluate model fairness and apply pre- and post-processing bias mitigation techniques. Ensuring diversity in training data and conducting regular audits is also essential.

**⚖️ Trade-offs (10 points)**

**Q: Discuss the trade-off between model interpretability and accuracy in healthcare.**  
In healthcare, **model interpretability** is critical for trust, compliance, and clinical validation. Doctors must understand why a model made a decision, especially when patient lives are at stake. While complex models like deep neural networks may offer higher accuracy, they often act as "black boxes." In contrast, simpler models like decision trees or logistic regression are more interpretable but might be less accurate.

**Q: If the hospital has limited computational resources, how might this impact model choice?**  
Limited resources constrain the use of compute-heavy models. Lightweight models such as **Random Forests**, **Gradient Boosted Trees**, or **Logistic Regression** may be preferred due to faster training and inference times. These models also simplify deployment and integration with hospital systems without requiring GPUs or cloud-based infrastructure.

**🧭 Part 4: Reflection & Workflow Diagram (10 Points)**

**💭 Reflection (5 points)**

**Q: What was the most challenging part of the workflow? Why?**  
The most challenging part was **data preprocessing**, especially handling missing values and balancing class distribution. Real-world datasets are often messy, and cleaning them without losing critical information is time-consuming. It also requires domain knowledge to engineer meaningful features without introducing bias.

**Q: How would you improve your approach with more time/resources?**  
With more time and resources, we would invest in **automated data profiling tools**, use **cloud-based notebooks** for scalability, and conduct **user testing** of model predictions. We’d also implement **model explainability tools** like SHAP to better understand and visualize prediction reasoning.

**📈 Workflow Diagram (5 points)**

Here's a basic AI Development Workflow diagram you can sketch or include using a tool like draw.io, Excalidraw, or Canva:

┌────────────────────┐

│ Problem Definition│

└────────┬───────────┘

↓

┌────────────────────┐

│ Data Collection │

└────────┬───────────┘

↓

┌────────────────────┐

│Data Preprocessing │

└────────┬───────────┘

↓

┌────────────────────┐

│ Model Selection │

└────────┬───────────┘

↓

┌────────────────────┐

│ Model Training │

└────────┬───────────┘

↓

┌────────────────────┐

│ Model Evaluation │

└────────┬───────────┘

↓

┌────────────────────┐

│ Model Deployment │

└────────────────────┘