Data Preparation & Preprocessing

1. **Scenario**: You're working with customer reviews in multiple languages, but your LSTM model was trained only on English text. What preprocessing steps would you add to handle multilingual input, and how would this affect your model's architecture?
2. **Scenario**: Your text dataset contains 30% missing values in the 'review_text' column. Propose three different strategies to handle these missing values and justify which one you would choose for LSTM classification.

Model Architecture & Training

3. **Scenario**: When training your LSTM on a 10GB dataset of product reviews, you notice the training time per epoch is 8 hours. Suggest three specific modifications to your model architecture or training process that could reduce this time while maintaining accuracy.
4. **Scenario**: Your LSTM model achieves 95% training accuracy but only 60% validation accuracy on sentiment analysis. Diagnose three potential causes and propose specific solutions for each.

Hyperparameter Tuning

5. **Scenario**: You need to deploy your LSTM model on mobile devices with limited memory. What three hyperparameters would you prioritize for optimization to reduce model size, and how would you adjust them?
6. **Scenario**: When testing different sequence lengths (50, 100, 200 tokens) for padding, you notice longer sequences decrease accuracy. Explain this phenomenon and propose two solutions.

Real-world Deployment

7. **Scenario**: Your deployed LSTM API receives a request containing emojis (e.g., "I love this product! 💔😊"). How would you modify your text preprocessing pipeline to handle this, and what impact might emojis have on classification?
8. **Scenario**: Users report that your sentiment analysis model misclassifies sarcastic comments (e.g., "Oh great, another delayed flight"). Propose two architectural changes to better detect sarcasm.

Performance Evaluation

9. **Scenario**: Your confusion matrix shows the model frequently confuses 'anger' (class 1) and 'fear' (class 3) in emotion detection. Design three targeted improvements to address this specific issue.
10. **Scenario**: When evaluating your LSTM on a new dataset from a different domain (e.g., trained on movie reviews, tested on medical forum posts), accuracy drops by 30%. Propose a transfer learning strategy to adapt the model.

Error Analysis & Debugging

11. **Scenario**: During inference, your LSTM outputs the same sentiment prediction regardless of input text. Identify three possible causes and the corresponding debugging steps.
12. **Scenario**: Gradients in your bidirectional LSTM explode during backpropagation. Suggest three specific modifications to stabilize training, explaining how each helps.

Advanced Applications

13. **Scenario**: You need to modify your text classification LSTM to simultaneously predict sentiment (positive/negative) and topic (sports/politics/tech). Design the modified architecture and loss function.
14. **Scenario**: Your LSTM must process documents exceeding 10,000 words while maintaining context. Propose two attention mechanisms you could integrate and explain their benefits.