**Case Related Questions**

1. **Case Introduction and Data Understanding:**

* What is the context of the credit risk prediction problem presented in the case?
* What are the key features in the dataset that are relevant for predicting loan approval?
* How can understanding the distribution of the target variable (loan approval) help in assessing the problem?

**2. Model Selection and Training:**

* Why were Logistic Regression, Decision Trees, Random Forest, and SVM chosen as potential algorithms for this classification problem?
* Explain the core concepts behind each of these algorithms in the context of classification.
* How was the dataset split into training and testing sets? Why is this important?

**3. Feature Scaling and Data Preprocessing:**

* Why did we need to perform feature scaling on the numeric features in the dataset?
* What is one-hot encoding, and why was it necessary for handling categorical variables like "Marital Status"?

**4. Model Evaluation and Comparison:**

* What metrics were used to evaluate the performance of the models? Why were these metrics chosen?
* Based on the evaluation results, which algorithm performed the best and why? Were there any trade-offs to consider?

**5. Model Interpretability:**

* Compare and contrast the interpretability of Decision Trees, Random Forest, Logistic Regression, and SVM. How do they differ in providing insights into model predictions?
* How can insights about feature importance be used to make decisions?

**6. Parameter Tuning and Generalization:**

* What is the purpose of parameter tuning in machine learning? Why is it important for model performance?
* Describe the process of grid search using cross-validation for finding the optimal parameters for each algorithm (bonus).
* How does parameter tuning impact model generalization? What are the risks of overfitting and underfitting?

**7. Business Implications and Ethical Considerations:**

* How could the insights from the best-performing model be used in a real-world banking scenario?
* What are potential ethical considerations when using machine learning to predict credit risk? How can bias be mitigated?

**8. Reflecting on Learning:**

* How has this case improved your understanding of machine learning algorithms and their application in real-world problems?
* What challenges did you encounter during the process of implementing and interpreting the models?
* How might you apply the knowledge gained from this case to other classification problems or real-world scenarios?