# Machine learning lab final exam

## Question: Build and Evaluate a Machine Learning Model

**Task:**

1. **Dataset Selection:**
   - Choose a publicly available dataset from sources such as Kaggle, UCI Machine Learning Repository, or another relevant source. Ensure the dataset is suitable for classification or regression tasks.

2. **Data Preprocessing:**
   - Load the dataset into your environment.
   - Perform exploratory data analysis (EDA). Include:
     - Summary statistics.
     - Visualizations (e.g., histograms, scatter plots, box plots).
   - Handle any missing values, encode categorical variables, and normalize/standardize features as necessary.

3. **Model Implementation:**
   - Select any three machine learning algorithms from the following options:
     - Logistic Regression
     - Decision Tree
     - Random Forest
     - Gradient Boosting (e.g., XGBoost)
   - Split the dataset into training and testing sets (e.g., 70/30 split).
   - Train your selected model on the training set.

4. **Model Evaluation:**
   - Evaluate the model using appropriate metrics for the chosen task:
     - For classification: Accuracy, Precision, Recall, F1-Score, Confusion Matrix.
     - For regression: Mean Absolute Error (MAE), Mean Squared Error (MSE), R-squared.
   - Provide visualizations of your evaluation metrics (e.g., ROC curve for classification).

5. **Hyperparameter Tuning:**
   - Implement hyperparameter tuning using techniques such as Grid Search or Random Search. Report the best parameters and their impact on model performance.

6. **Conclusion:**
   - Summarize your findings. Discuss:
     - The performance of the model based on evaluation metrics.
     - Any challenges faced during the analysis.
     - Recommendations for improving the model.

7. **Code Submission:**
   - Submit your complete code along with a report that documents each step of your process, including explanations of your choices and findings.