**Phase 2: Implementation of Machine Learning Models and Hyperparameter Tuning**

**Problem Statement:**

In Phase 1 of the project, we preprocessed the Titanic dataset and prepared it for machine learning modeling. Now, in Phase 2, our objective is to implement different machine learning models on the dataset and optimize their hyperparameters to improve performance.

**Approach:**

**1. Data Preprocessing:**

- Handled missing values for 'Age' and 'Fare' columns by imputing the median values.

- Dropped the 'Cabin' column due to too many missing values.

- Engineered the 'FamilySize' feature by combining 'SibSp' and 'Parch' columns.

- Encoded categorical variables ('Sex' and 'Embarked') using one-hot encoding.

**2. Feature Scaling:**

- Standardized the numerical features 'Age' and 'Fare' using StandardScaler.

**3. Model Implementation:**

- Implemented four machine learning models:

- Logistic Regression

- Random Forest

- Support Vector Machine (SVM)

- Gradient Boosting

- Each model was trained and evaluated on the training dataset.

**4. Hyperparameter Tuning:**

- Tuned hyperparameters for each model to find the optimal settings.

- Explored multiple values of learning rates and epochs for Gradient Boosting.

- Used a grid search approach to find the best combination of hyperparameters.

**5. Evaluation:**

- Evaluated model performance using training and validation accuracy scores.

- Plotted change in error over epochs for each model to visualize convergence.

- Created a table summarizing results for each model and highlighting the best performing configuration.

**Results:**

- Random Forest emerged as the best performing model, exhibiting high accuracy and robustness to overfitting.

- Gradient Boosting also showed competitive performance, especially with optimal hyperparameter settings.

- The visualization of error convergence over epochs provided insights into model training dynamics and helped in selecting appropriate hyperparameters.

**Conclusion:**

In Phase 2, we successfully implemented and optimized machine learning models on the Titanic dataset. The thorough exploration of hyperparameters and visualization of model training dynamics enhanced our understanding of model behavior and facilitated informed decision-making. Moving forward, we will proceed with model deployment and further analysis of model predictions.