LGBM & XGBM

**Objective:**

The objective of this assignment is to compare the performance of Light GBM and XG Boost algorithms using the Titanic dataset.

**Exploratory Data Analysis (EDA):**

1. Load the Titanic dataset using Python's pandas library.
2. Check for missing values.
3. Explore data distributions using histograms and box plots.
4. Visualize relationships between features and survival using scatter plots and bar plots.

**Data Preprocessing:**

1. Impute missing values.
2. Encode categorical variables using one-hot encoding or label encoding.
3. If needed you can apply more preprocessing methods on the given dataset.

**Building Predictive Models:**

1. Split the preprocessed dataset into training and testing sets.
2. Choose appropriate evaluation metrics (e.g., accuracy, precision, recall, F1-score) for model evaluation.
3. Build predictive models using LightGBM and XGBoost algorithms.
4. Train the models on the training set and evaluate their performance on the testing set.
5. Use techniques like cross-validation and hyperparameter tuning to optimize model performance.

**Comparative Analysis:**

1. Compare the performance metrics (e.g., accuracy, precision, recall) of LightGBM and XGBoost models.
2. Visualize and interpret the results to identify the strengths and weaknesses of each algorithm.

**Submission Requirements:**

Well-commented code explaining each step of the analysis.

Visualizations with appropriate titles and labels.

A brief report summarizing the comparative analysis results and practical implications.

**Answer:**

<https://colab.research.google.com/drive/1_zAKKbPWLVmuCVeJYsNfYb9hycY3-q8L?usp=sharing>