AI Lab Test Report

(GROUP H)

This report summarizes the AI Lab Test analysis performed on the **StudentPerformance.csv** dataset. The workflow includes feature engineering, model training, hyperparameter tuning, explainability with SHAP, and visual interpretation using regression-based evaluation metrics.

Task	Description
1. Feature Engineering	Created new metrics like distance_to_goal, performance_angle, and is_high_efficiency
2. Data Handling	Imputed missing numeric data using median strategy and encoded categorical columns
3. Model Training	Trained Linear Regression, Random Forest, and Gradient Boosting models fo performance
4. Hyperparameter Tuning	Optimized RandomForestRegressor using GridSearchCV with R^2 scoring.
5. Explainability	Used SHAP to visualize feature contributions.
6. Visualization	Generated feature importance, predicted vs actual, and residual distribution plots.

The Random Forest model achieved the best predictive performance among tested models. Evaluation metrics were based on regression analysis (R^2, MAE, and MSE). Cross-validation confirmed consistent model reliability. Feature importance analysis showed that **Hours Studied**, **Previous Scores**, and **Sleep Hours** were the top predictors of student performance.

The AI Lab Test successfully demonstrated data preprocessing, feature engineering, model comparison, and explainability. This process provides insights into factors influencing student academic performance and highlights reproducible ML workflow design.