

Regression Model Performance Analysis

This document outlines the evaluation and comparison of several regression models trained to predict student scores. The aim was to identify the most effective model based on its predictive capabilities on unseen data.

Model Comparison Results

Following the training of four distinct regression models—Linear Regression, Decision Tree Regressor, Random Forest Regressor, and Support Vector Regression—their performance was meticulously compared. The primary criterion for ranking these models was their **Test R² scores**. This metric is crucial as it quantifies how well each model explains the variance in student scores when presented with new, unobserved data. The comparison highlighted clear differences in performance across the algorithms.

Performance Interpretation

To provide a consistent and understandable interpretation of the models' R² scores, a clear classification system was established:

- **Excellent:** $R^2 > 0.8$ (The model explains over 80% of the variance in scores).
- **Good:** $R^2 > 0.6$ (The model explains between 60% and 80% of the variance).
- **Fair:** $R^2 > 0.4$ (The model explains between 40% and 60% of the variance).
- **Poor:** $R^2 \leq 0.4$ (The model explains less than 40% of the variance).

Evaluation Methodology

The overall evaluation process adopted a comprehensive approach, combining multiple analytical techniques:

- **Quantitative Assessment:** This involved the rigorous use of key regression metrics:
 - **R² (Coefficient of Determination):** To provide a numerical measure of the proportion of variance in the dependent variable that is predictable from the independent variables.
 - **MAE (Mean Absolute Error):** To indicate the average magnitude of the errors in a set of predictions, without considering their direction.
 - **RMSE (Root Mean Squared Error):** To measure the square root of the average of the squared errors, giving a relatively high weight to large errors.

- **Visual Analysis:** Scatter plots depicting actual versus predicted values were generated. Additionally, residual plots were created to visually inspect for any discernible patterns in prediction errors or the presence of outliers, which could indicate model biases or limitations.
- **Comparative Analysis:** A side-by-side comparison of all implemented models was conducted. This objective analysis allowed for the direct determination of the optimal performer based on their demonstrated predictive power across the established metrics.

The final selection of the "best model" was an automated process, driven strictly by the model that achieved the **highest Test R^2 score**, ensuring an unbiased choice grounded in empirical performance.