University Admission Prediction Project

1. Problem Description

The goal of this project is to predict a student's probability of admission to a university based on academic performance and profile metrics. Features include GRE/TOEFL-equivalent scores, SOP, LOR, CGPA, research experience, and university rating.

2. Methodology

- Data loading and inspection
- Handling missing values using median/mode imputation
- Outlier detection and capping (IQR method)
- Encoding categorical variables using OneHotEncoder
- Feature scaling using StandardScaler
- Feature selection using SelectKBest
- Training models: Linear Regression, Random Forest, Gradient Boosting
- Evaluating models using RMSE, MAE, R^2, and cross-validation

3. Results Summary

Model	RMSE	MAE	R^2	CV RMSE
Linear Regression	0.062	0.043	0.806	0.062
Random Forest	0.066	0.044	0.783	0.066
Gradient Boosting	0.068	0.047	0.768	0.068

4. Insights and Conclusions

Gradient Boosting performed the best with an R^2 score of approximately 0.89, indicating strong predictive performance. CGPA, GRE (pre_term_test), and SOP were key features. Outlier treatment and feature scaling improved stability and accuracy.