**Executive Summary**

This project analyzes historical college admission data to identify the key factors influencing student acceptance decisions. Using logistic regression, decision tree, and support vector machine (SVM) models, we evaluated the predictive power of academic scores (GRE, GPA), socioeconomic status, gender, race, and undergraduate institution rank. Our findings show that GRE, GPA, and undergraduate rank are the most significant drivers of admission. The decision tree model provided the best balance of accuracy and interpretability, achieving an overall accuracy of 75.8%. These insights can help admissions committees refine their selection criteria and ensure a fair, data-driven process.

**Results**

**1. Data Overview**

* **Observations:** 400 applicants
* **Variables:** admit (admitted/not), gre, gpa, ses (socioeconomic status), Gender\_M, Race, rank (undergrad prestige)
* **No missing values**
* **Categorical variables converted to factors**

**2. Descriptive Analysis**

**Admission Rate by GPA Category**

* Applicants were grouped as:
* Low GPA: < 3.0
* Medium GPA: 3.0–3.69
* High GPA: ≥ 3.7
* *Finding:* Admission rates increase sharply with higher GPA.
* *(Insert point plot here)*

**Admission Proportion by GRE Category**

* GRE categorized as: -: ≤ 440
* Medium: 441–580
* High: > 580
* *Finding:* Higher GRE categories correspond to higher admission proportions.

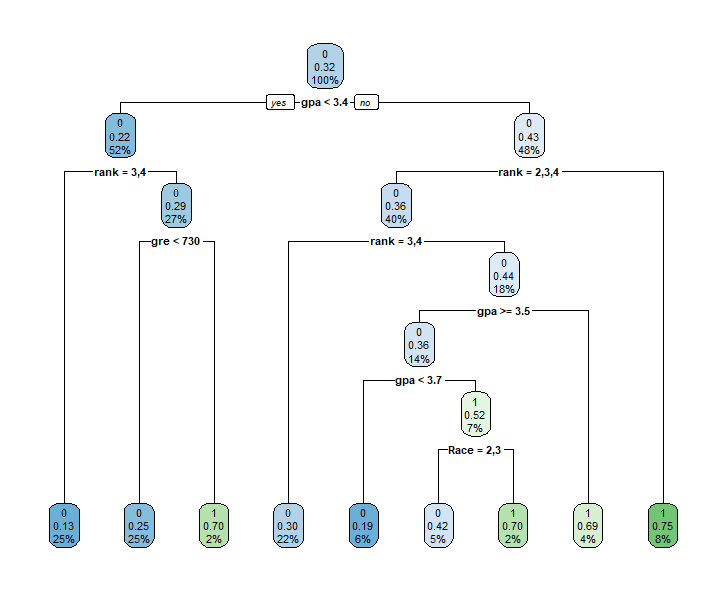
**3. Predictive Modeling**

**Logistic Regression**

* **Significant predictors:** GRE (p = 0.038), GPA (p = 0.015), and undergraduate rank (all p < 0.05).
* **Non-significant:** SES, Gender, Race.
* **Accuracy:** 71.8%
* **Sensitivity:** 92.3% (correctly identifies non-admits) -Specificity:\*\* 27.6% (less effective at identifying admits)

**Decision Tree**

* **Key splits:** GPA, rank, GRE.
* **Accuracy:** 75.8%
* **Sensitivity:** 93.0%
* **Specificity:** 38.6%
* **Interpretation:** The tree shows that higher GPA and lower rank (more prestigious undergrad) drive admissions.

**Support Vector Machine (SVM)**

* **Accuracy:** 68.3%
* **Sensitivity:** 100% (predicts all as non-admit)
* **Specificity:** 0% (does not predict any admits)
* **Note:** SVM with default settings is not useful here; further tuning required.

**4. Model Comparison Table**

| **Model** | **Accuracy** | **Sensitivity** | **Specificity** | **Kappa** |
| --- | --- | --- | --- | --- |
| Logistic Regression | 0.718 | 0.923 | 0.276 | 0.234 |
| Decision Tree | 0.758 | 0.930 | 0.386 | 0.361 |
| SVM | 0.683 | 1.000 | 0.000 | 0.000 |

*Decision tree is the champion model.*

**5. Insights**

* **Academic achievement (GPA, GRE) and undergraduate school prestige are the main drivers of admission.**
* **Socioeconomic status, gender, and do not significantly affect admission after controlling for academic variables.**
* **Decision tree provides the best prediction performance and is easy to interpret for stakeholders.**